

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 04:01:44 ; Search time 25622 Seconds

11676.376 Million cell updates/sec
 (without alignments)Title: US-09-242-772-116
 Perfect score: 7313

Sequence: 1 ggccgcatacacataat.....tatgaataaaatctgtgcc 7313

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters:
 Minimum DB seq length: 0
 Maximum DB seq length: 20000000000Post-processing: Minimum Match 0%
 Maximum Match 100%Listing first 45 summaries
 Database : GenBmbl1:
 1: gb_ba;*: 2: gb_lntg;*: 3: gb_in;*: 4: gb_on;*: 5: gb_ov;*: 6: gb_dat;*: 7: gb_ph;*: 8: gb_dl;*: 9: gb_dx;*: 10: gb_ro;*: 11: gb_sts;*: 12: gb_sy;*: 13: gb_un;*: 14: gb_vl;*: 15: em_da;*: 16: em_fun;*: 17: em_hum;*: 18: em_in;*: 19: em_mu;*: 20: em_om;*: 21: em_xr;*: 22: em_ov;*: 23: em_dat;*: 24: em_ph;*: 25: em_pl;*: 26: em_ro;*: 27: em_sts;*: 28: em_un;*: 29: em_vl;*: 30: em_htg_hum;*: 31: em_htg_inv;*: 32: em_htg_other;*: 33: em_htg_mus;*: 34: em_htg_pln;*: 35: em_htg_rod;*: 36: em_htg_mam;*: 37: em_htg_vrt;*: 38: em_sy;*: 39: em_htg_hum;*: 40: em_htg_mus;*: 41: em_htg_other;*

score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	7313	100.0	7313	9 HSU65002	Human zinc finger protein Plagi mRNA, complete cds.
2	7312.6	100.0	7313	6 A9511	Human zinc finger protein Plagi mRNA, complete cds.
3	6520	89.2	142102	9 AC107952	Human zinc finger protein Plagi mRNA, complete cds.
4	2821.6	38.6	182102	2 AC097274	Mus musculus zinc finger protein Plagi mRNA, complete cds.
5	2821.6	38.6	241148	10 AL8079387	Mouse DNA sequence
6	2761.2	37.8	278377	2 AC1298939	Rattus norvegicus zinc finger protein Plagi mRNA, complete cds.
7	2623.8	35.9	249552	2 AC123210	Rattus norvegicus zinc finger protein Plagi mRNA, complete cds.
8	1295.6	17.7	1575	10 AF057166	Human zinc finger protein Plagi mRNA, complete cds.
9	707.4	9.7	65748	2 AC090187	Human zinc finger protein Plagi mRNA, complete cds.
10	651.8	8.9	65748	2 AC090187	Human zinc finger protein Plagi mRNA, complete cds.
11	645.4	8.8	49357	2 AC107961	Human zinc finger protein Plagi mRNA, complete cds.
12	416	5.7	421	11 G25280	Human zinc finger protein Plagi mRNA, complete cds.
13	392.8	5.4	1829	9 BC023655	Human zinc finger protein Plagi mRNA, complete cds.
14	392.8	5.4	2429	9 AK026936	Human zinc finger protein Plagi mRNA, complete cds.
15	392.8	5.4	4051	9 AF006005	Human zinc finger protein Plagi mRNA, complete cds.
16	392.8	5.4	5638	9 D83784	Human zinc finger protein Plagi mRNA, complete cds.
17	391.2	5.3	5267	10 AB051854	Mus musculus zinc finger protein Plagi mRNA, complete cds.
18	383.8	5.2	550	9 AF221550	Human zinc finger protein Plagi mRNA, complete cds.
19	364.8	5.0	1437	10 AF181262	Human zinc finger protein Plagi mRNA, complete cds.
20	359.2	4.9	161478	9 AC025538	Human zinc finger protein Plagi mRNA, complete cds.
21	352.8	4.8	145414	9 HSA392M18	Human zinc finger protein Plagi mRNA, complete cds.
22	352.6	4.8	200329	10 AC078911	Human zinc finger protein Plagi mRNA, complete cds.
23	352.6	4.8	208614	10 AL807380	Human zinc finger protein Plagi mRNA, complete cds.
24	352	4.8	240510	2 AC134371	Rattus norvegicus zinc finger protein Plagi mRNA, complete cds.
25	351.6	4.8	2334	6 BD005470	Nucleic acid sequence
26	351.6	4.8	2828	9 HSA635	Human zinc finger protein Plagi mRNA, complete cds.
27	351.6	4.8	3215	9 AK096531	Human zinc finger protein Plagi mRNA, complete cds.
28	351.6	4.8	3663	9 HSA311395	Human zinc finger protein Plagi mRNA, complete cds.
29	351	4.8	240336	2 AC094844	Rattus norvegicus zinc finger protein Plagi mRNA, complete cds.
30	350	4.8	2815	6 AX329760	Sequence
31	350	4.8	3162	9 AK091707	Human zinc finger protein Plagi mRNA, complete cds.
32	350	4.8	4833	9 HSU7621	Human zinc finger protein Plagi mRNA, complete cds.
33	344.4	4.7	1874	5 AF186476	Human zinc finger protein Plagi mRNA, complete cds.
34	333	4.6	376	6 AR274991	Sequence
35	301.8	4.1	5028	10 RN072620	Rattus norvegicus zinc finger protein Plagi mRNA, complete cds.
36	295.8	4.0	2350	10 AF324471	Mus musculus zinc finger protein Plagi mRNA, complete cds.
37	294.4	4.0	2771	10 MMZINC_FPR	Zinc finger protein Plagi mRNA, complete cds.
38	294.4	4.0	2790	6 BD005460	Zinc finger protein Plagi mRNA, complete cds.
39	294.4	4.0	3732	10 MMZINC_FIP	Zinc finger protein Plagi mRNA, complete cds.
40	294.4	4.0	3975	10 AF147785	Zinc finger protein Plagi mRNA, complete cds.
41	291.4	4.0	236442	2 AC128089	Zinc finger protein Plagi mRNA, complete cds.
42	279.2	3.8	2467	9 HSA303119	Zinc finger protein Plagi mRNA, complete cds.
43	279.2	3.8	137397	9 HSU468K18	Zinc finger protein Plagi mRNA, complete cds.
44	277.6	3.8	2561	6 A69513	Zinc finger protein Plagi mRNA, complete cds.
45	277.6	3.8	2561	9 HSU81992	Zinc finger protein Plagi mRNA, complete cds.

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
1	HSU65002	Human zinc finger protein Plagi mRNA, complete cds.	U65002	PRI 14-FEB-1997			
		1 (bases 1 to 7313)					
		Kas K., Voz,M.L., Roijer,E., Astrom,A.K., Meyen,E., Stenman,G. and Van de Ven,W.J.					
		Promoter swapping between the genes for a novel zinc finger protein					

Pred. No. is the number of results predicted by chance to have a

and beta-catenin in pleiomorphic adenomas with $t(3;8)$ (p21;q12) translocations

QY	3661	AACAGAAGAAATATTCTGAGCACTAATTAGCCCTATATGGAAACTTCTTCCTT	3720	Db	4741	AAATTGACTGTAGTGGTCACTAACCCCTTTGAGGAATTCTATTAAGATGGCAG 4800
Dy	3661	AAACAGAAGAAATATTCTGAGCACTAATTAGCCCTATATGGAAACTTCTTCCTT	3720	Db	4741	AAATTGACTGTAGTGGTCACTAACCCCTTTGAGGAATTCTATTAAGATGGCAG 4800
QY	3721	TTCGAGGGCAAGTTCACTGCAGATGGTCACTGCAGATGGTCACTGCAGATGGT	3780	Db	4801	ACTCGCTTATTGATGTTCACTGCAGATGGTCACTGCAGATGGTCACTGCAGATGGT
Dy	3721	TTCGAGGGCAAGTTCACTGCAGATGGTCACTGCAGATGGTCACTGCAGATGGT	3780	Db	4801	ACTCGCTTATTGATGTTCACTGCAGATGGTCACTGCAGATGGTCACTGCAGATGGT
QY	3781	GCGATTACTGATAATACTAATTAAATTTGGGGTAGTATAATTGGTGGAGA	3840	Db	4861	TGCCCTAGAAAAGATGTTGTTCTATGTCAGTAGATGTTAACAGAATGGCTTT
Dy	3781	GCGATTACTGATAATACTAATTAAATTTGGGGTAGTATAATTGGTGGAGA	3840	Db	4861	TGCCCTAGAAAAGATGTTGTTCTATGTCAGTAGATGTTAACAGAATGGCTTT
QY	3841	ATGTAATGTTAATATGTAATGGGGGGTTCATATATAACACACAC	3900	Db	4861	TGCCCTAGAAAAGATGTTGTTCTATGTCAGTAGATGTTAACAGAATGGCTTT
Dy	3841	ATGTAATGTTAATATGTAATGGGGGGTTCATATATAACACACAC	3900	Db	4861	TGCCCTAGAAAAGATGTTGTTCTATGTCAGTAGATGTTAACAGAATGGCTTT
QY	3901	ACACATGCAACACATACGGACTCATAGATAAACAGTGGCTTG	3960	Db	4981	CCTAGCTTGGTTCATCTCTAGGCCGAACTATATCCTGAAATTCTGTCAGT
Dy	3901	ACACATGCAACACATACGGACTCATAGATAAACAGTGGCTTG	3960	Db	4981	CCTAGCTTGGTTCATCTCTAGGCCGAACTATATCCTGAAATTCTGTCAGT
QY	3961	CTCCCTGATATTATCATGCTCTATATTTCATCTGGAGCTATAC	4020	Db	4981	CCTAGCTTGGTTCATCTCTAGGCCGAACTCTAGTTCACTTCACATGTCAGT
Dy	3961	CTCCCTGATATTATCATGCTCTATATTTCATCTGGAGCTATAC	4020	Db	4981	CCTAGCTTGGTTCATCTCTAGGCCGAACTCTAGTTCACTTCACATGTCAGT
QY	4021	TATGTTAGATTAAATTATATGAAATTACATTATGAAAGGTC	4080	Db	5041	TAGTTAGATTGGCCCTAGACAGATAATTGGGTTCACTTCACATGTCAGT
Dy	4021	TATGTTAGATTAAATTATGAAAGGTC	4080	Db	5041	TAGTTAGATTGGCCCTAGACAGATAATTGGGTTCACTTCACATGTCAGT
QY	4081	ATTAATTATATATGTTAACTTCTGAGCTATCTGTTAGT	4140	Db	5101	TATGTAACAGTCTACATGTTAGTGGAAAGGTCTTGGCACCTAA
Dy	4081	ATTAATTATATATGTTAACTTCTGAGCTATCTGTTAGT	4140	Db	5101	TATGTAACAGTCTACATGTTAGTGGAAAGGTCTTGGCACCTAA
QY	4141	TTTGTGTTAGGAGAAATTCTGAAAGTAGGCTTTGCTCTTACTGTT	4200	Db	5161	GCAATGCTATCCATTGATGTTACAGTACCTTCATGTTAAACTT
Dy	4141	TTTGTGTTAGGAGAAATTCTGAAAGTAGGCTTTGCTCTTACTGTT	4200	Db	5161	GCAATGCTATCCATTGATGTTACAGTACCTTCATGTTAAACTT
QY	4201	CCTTGTGTTCTATGAAAGTAGGCTTTGCTCTTACTGTTAGTGTCA	4260	Db	5221	TCAATTAAATACCTTAACTTACAGTCACTTCTTAACTGTTAACTT
Dy	4201	CCTTGTGTTCTATGAAAGTAGGCTTTGCTCTTACTGTTAGTGTCA	4260	Db	5221	TCAATTAAATACCTTAACTTACAGTCACTTCTTAACTGTTAACTT
QY	4261	TTATTGCTTAACTTACCATATAAGGAAATTGGTCACTTCTGTTAAA	4320	Db	5281	ATGCACTACTCTGTTGCACTTTTATTCACTTGTGTTCTGTT
Dy	4261	TTATTGCTTAACTTACCATATAAGGAAATTGGTCACTTCTGTTAAA	4320	Db	5281	ATGCACTACTCTGTTGCACTTTTATTCACTTGTGTTCTGTT
QY	4321	CTATTTTAAATAGAAATTCAAATAGCTGTTGAACTTCTGTTAAA	4380	Db	5341	TTTGCCTATTGTTGCTTCTTACTGTTCTTACTGTTCTGTT
Dy	4321	CTATTTTAAATAGAAATTCAAATAGCTGTTGAACTTCTGTTAAA	4380	Db	5341	TTTGCCTATTGTTGCTTCTTACTGTTCTTACTGTTCTGTT
QY	4381	ATGCTACTAAGAGTCATCAATTAGTGAATCTGGCTTACTGGCTGTT	4440	Db	5401	TTGTGTTAATTGGCAAGTAGGCTTAACTTAACTGTTGCTT
Dy	4381	ATGCTACTAAGAGTCATCAATTAGTGAATCTGGCTTACTGGCTGTT	4440	Db	5401	TTGTGTTAATTGGCAAGTAGGCTTAACTTAACTGTTGCTT
QY	4441	ACTGCTATTAAACGGTACCCAGATTATTAGTAAGCTTGGCTTACTGGCTGTT	4500	Db	5461	GAGCCATTGGCTTAACTTAACTGTTGCTTAACTGTTGCTT
Dy	4441	ACTGCTATTAAACGGTACCCAGATTATTAGTAAGCTTGGCTTACTGGCTGTT	4500	Db	5461	GAGCCATTGGCTTAACTTAACTGTTGCTTAACTGTTGCTT
QY	4501	ATAGGCCCATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	4560	Db	5521	TCATCTAGTGGCAATTACACATTACAGATATGCTTGGCTT
Dy	4501	ATAGGCCCATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	4560	Db	5521	TCATCTAGTGGCAATTACACATTACAGATATGCTTGGCTT
QY	4561	ACACATCATCTAACTGGCTGTTGTTGTTGTTGTTGTTGTTGTT	4620	Db	5581	CAACCTCTATACATACATACAGATATGCTTGGCTT
Dy	4561	ACACATCATCTAACTGGCTGTTGTTGTTGTTGTTGTTGTTGTT	4620	Db	5581	CAACCTCTATACATACAGATATGCTTGGCTT
QY	4621	GGCACCGTGTAAATTGGCTTCACTTCACTGGCTTCACTGGCTT	4680	Db	5701	TATTTAAACAGGTCACTTCACTGGCTTCACTGGCTTCACTGGCTT
Dy	4621	GGCACCGTGTAAATTGGCTTCACTTCACTGGCTTCACTGGCTT	4680	Db	5701	TATTTAAACAGGTCACTTCACTGGCTTCACTGGCTTCACTGGCTT
QY	4681	CCACCCCCAGACTGTTAACTTCACTGGCTTCACTGGCTTCACTGGCTT	4740	Db	5761	TATTTAAACAGGTCACTTCACTGGCTTCACTGGCTTCACTGGCTT
Dy	4681	CCACCCCCAGACTGTTAACTTCACTGGCTTCACTGGCTTCACTGGCTT	4740	Db	5761	TATTTAAACAGGTCACTTCACTGGCTTCACTGGCTTCACTGGCTT

Qy	61	GCGCGTCAAGCCGAATAATGAAAAATTATGAAATTTCGGGGCGTGTAGAG	120		1141	AGATTGGGAAAGGATCACCTGACTGACATGAGAGTACAATGAGACCTT	1200
Db	61	GCGCGTCAAGCCGAATAATGAAAAATTATGAAATTTCGGGGCGTGTAGAG	120		1201	CITGAAGTCAAACAGAACGACTGGATTCCCTGACCCATTACCTCCTAATG	1260
Qy	121	GCGCGTCAAGCCGAATAAGCCCCGGTAAAGCTCTGGCTGCTTGCTT	180		1201	CITGAAGTCAAACAGAACGACTGGATTCCCTGACCCATTACCTCCTAATG	1260
Db	121	GCGCGTCAAGCCGAATAAGCCCCGGTAAAGCTCTGGCTGCTTGCTT	180		1261	CTTATAAAGGAGCTCCCTCCGGTAGTCCTTCCAGTCAACTCCATTG	1320
Qy	181	GCGCGTCAAGCCGAATAAGCTCTGGACCCGAACTGGCTTCTTAATAAG	240		1261	CTTATAAAGGAGCTCCCTCCGGTAGTCCTTCCAGTCAACTCCATTG	1320
Db	181	GCGCGTCAAGCCGAATAAGCTCTGGACCCGAACTGGCTTCTTAATAAG	240		1261	CTTATAAAGGAGCTCCCTCCGGTAGTCCTTCCAGTCAACTCCATTG	1320
Qy	241	CTAGAGTTGGAGAGGGCTCAGATGCCAAATGGAAAGGATTGAACTCT	300		1321	CCATCAAAACAACTTGCAGTTAACCTCTAACACCTCCATTGAGAGC	1380
Db	241	CTAGAGTTGGAGAGGGCTCAGATGCCAAATGGAAAGGATTGAACTCT	300		1321	CCATCAAAACAACTTGCAGTTAACCTCTAACACCTCCATTGAGAGC	1380
Db					1381	TCCGGATCTGCCACCAAAATGATGACACACTTACCTTGGGAAATG	1440
Qy					1381	TCCGGATCTGCCACCAAAATGATGACACACTTACCTTGGGAAATG	1440
Db					1441	ATGGACACTGTTCATCCCTCTCACACACCTTCTCAATATGGTCA	1500
Qy					1441	ATGGACACTGTTCATCCCTCTCACACACCTTCTCAATATGGTCA	1500
Db	301	TCCACGAAAGTCATGGACATGGCTCAAGTCAATGAACTTCCATG	360		1441	ATGGACACTGTTCATCCCTCTCACACACCTTCTCAATATGGTCA	1500
Db	301	TCCACGAAAGTCATGGACATGGCTCAAGTCAATGAACTTCCATG	360		1441	ATGGACACTGTTCATCCCTCTCACACACCTTCTCAATATGGTCA	1500
Qy					1501	TATGCAATTCTCATCCCTGAAAGGAAATTGAGAGTTACCTG	1560
Db					1501	TATGCAATTCTCATCCCTGAAAGGAAATTGAGAGTTACCTG	1560
Qy					1561	ATGGAGTTACAAGGCTGCCTCTTCTCTGAAAGAACGCCCATA	1620
Db					1561	ATGGAGTTACAAGGCTGCCTCTTCTCTGAAAGAACGCCCATA	1620
Qy					1561	ATGGAGTTACAAGGCTGCCTCTTCTCTGAAAGAACGCCCATA	1620
Db					1621	AAGCTTGGGTTGGATCCCTAGATGGTCCCTAGATGATGTTGAG	1680
Qy					1621	AAGCTTGGGTTGGATCCCTAGATGGTCCCTAGATGATGTTGAG	1680
Db					1681	TCCAAAAGCTCTACTCCATGAGCTTCCAAACTGTCGAC	1740
Qy					1681	TCCAAAAGCTCTACTCCATGAGCTTCCAAACTGTCGAC	1740
Db					1741	TTGTTTAATTTCATACCCTTAATCCCTCATGTTCCACCTCCC	1800
Qy					1741	TTGTTTAATTTCATACCCTTAATCCCTCATGTTCCACCTCCC	1800
Db					1801	GGATGAGCTTCCAGGAGAACGACATCTCTGTTCCACCTCCC	1860
Qy					1801	GGATGAGCTTCCAGGAGAACGACATCTCTGTTCCACCTCCC	1860
Db					1801	GGATGAGCTTCCAGGAGAACGACATCTCTGTTCCACCTCCC	1860
Qy					1861	CAGGATCTCAGGATCTCTCCATGAACTTCTGCTCTGACTCT	1920
Db					1861	CAGGATCTCAGGATCTCTCCATGAACTTCTGCTCTGACTCT	1920
Qy					1981	TAGGATTCTGGACATGGATTCAATGAAATGTGTTGCTTAATG	2040
Db					1981	TAGGATTCTGGACATGGATTCAATGAAATGTGTTGCTTAATG	2040
Qy					2101	ATTTCATTAACTTGTGCTACTTAAACCTATAAAATTCCTG	2100
Db					2101	ATTTCATTAACTTGTGCTACTTAAACCTATAAAATTCCTG	2100
Qy					2161	GTGTCAGTTAACTTGTGCTACTTAAACCTATAAAATTCCTG	2220
Db					2161	GTGTCAGTTAACTTGTGCTACTTAAACCTATAAAATTCCTG	2220
Qy					2221	AGATTTTTTTTCCATGAAAGCCATTATTAGTAGAACTTCCCA	2280
Db					2221	AGATTTTTTTTCCATGAAAGCCATTATTAGTAGAACTTCCCA	2280
Qy							

Db	2221	AGACCTTTTTCATATAGAAAGCCATTATAGTAGTAACCTTTACAATCCCA	2280	QY	3361	ATAAGTTGAGTTCTGTAGGGTAAGAACAAAGTAAAGTAAACCTATCTTGTTAACAT	3420
Db	2281	TTTCAAACTACTTTAGATCTTAAATTTCATTTGCTTAAPAACTGGCTCTAAC	2340	Db	3361	ATAAGTTGAGTTCTGTAGGGTAAGAACAAAGTAAAGTAAACCTATCTTGTTAACAT	3420
QY	2341	TTTCAAACTACTTTAGCTTAAATTTCATTTGCTTAAPAACTGGCTCTAAC	2340	QY	3421	GAAAATAACATTGGAATATATTATCAGATAATTATAGCTTTCTTCACATT	3480
Db	2341	TTTCAAACTACTTTAGCTTAAATTTCATTTGCTTAAPAACTGGCTCTAAC	2400	Db	3421	GAAAATAACATTGGAATATATTATCAGATAATTATAGCTTTCTTCACATT	3480
QY	2401	TGAATAACTCAAGGGTTAAATTCTTACTAGCTCTAAATGCTTAATCAAGTC	2460	QY	3481	AGACACCACATATCTCTTAAGGTCTTAATTATATTCTAAGGTCAAGTGG	3540
Db	2401	TGAATAACTCAAGGGTTAAATTCTTACTAGCTCTAAATGCTTAATCAAGTC	2460	Db	3481	AGACACCACATATCTCTTAAGGTCTTAATTATATTCTAAGGTCAAGTGG	3540
QY	2461	TTCAATGAACTAACAGTCAGTTGGAGATAATAATGTTCTTAGATAACCCATA	2520	QY	3541	GACATAACCTAACAGAACATATCTTTAGAACCTTATTTAGAACCCCT	3600
Db	2461	TTCAATGAACTAACAGTCAGTTGGAGATAATAATGTTCTTAGATAACCCATA	2520	Db	3541	GACATAACCTAACAGAACATATCTTTAGAACCTTATTTAGAACCCCT	3600
QY	2521	TTTCAGATCAGTATTCGAAGACTCTGTCTGCTAAATATGCTCATCTTA	2580	QY	3601	TCCGTGCGGATGGGGAGGGGCCAGGTGGAGGTGTTAGAGAGTATATACTCTTTAA	3660
Db	2521	TTTCAGATCAGTATTCGAAGACTCTGTCTGCTAAATATGCTCATCTTA	2580	Db	3601	TCCCTGGGGATGGGGAGGGCCAGGTGGAGGTGTTAGAGAGTATATACTCTTTAA	3660
QY	2581	TTATGAGCCTTTAGGAALACAAACCCCTAACACAAGGATCACATTATGAAAG	2640	QY	3661	AAACAGAGAGAAATAATATTCTGAAACTCTTGCCTTATGAAACTCTTCCCT	3720
Db	2581	TTATGAGCCTTTAGGAALACAAACCCCTAACACAAGGATCACATTATGAAAG	2640	Db	3661	AAACAGAGAGAAATAATATTCTGAAACTCTTGCCTTATGAAACTCTTCCCT	3720
QY	2641	AGACTCTGTAGGTGACATGGCATTTGGTCACTTAATAGTGGCCCTAAATTAGTACA	2700	QY	3721	TTCTAGGCCAGTTATCACTGCAGATGCAAGTTCACCTTACCAAGAATTCTTAAATGAGT	3780
Db	2641	AGACTCTGTAGGTGACATGGCATTTGGTCACTTAATAGTGGCCCTAAATTAGTACA	2700	Db	3721	TTCTAGGCCAGTTATCACTGCAGATGCAAGTTCACCTTACCAAGAATTCTTAAATGAGT	3780
QY	2701	CAGGATTTTGTGTGTTCATCTCTTACATGCTCATCTTCAATTAAATAGTA	2760	QY	3781	GCAGATTACTGAAATAATACATTAAATAATTTGAGTAGTATAATTGTTGTTGAGA	3840
Db	2701	CAGGATTTTGTGTGTTCATCTCTTACATGCTCATCTTCAATTAAATAGTA	2760	Db	3781	GCAGATTACTGAAATAATACATTAAATAATTTGAGTAGTATAATTGTTGTTGAGA	3840
QY	2761	ATAGTGATGGCATTGGGCTCTCAGACTGATPATAGTAGATCTTGTCTTTC	2820	QY	3841	AATGTAATTGTAAATGTAATGGGGCTTCAATAATACACACAC	3900
Db	2761	ATAGTGATGGCATTGGGCTCTCAGACTGATPATAGTAGATCTTGTCTTTC	2820	Db	3841	AATGTAATTGTAAATGTAATGGGGCTTCAATAATACACACAC	3900
QY	2821	CACCTTCACATCAAGGGTGGTCAGTGCCAGCAATTATTTCATTTGTTGCC	2880	QY	3901	ACACATGCAACACATACCGCACTTCATAGATAACAAAGTTGCTCTGAAAGGGCTTGC	3960
Db	2821	CACCTTCACATCAAGGGTGGTCAGTGCCAGCAATTATTTCATTTGTTGCC	2880	Db	3901	ACACATGCAACACATACCGCACTTCATAGATAACAAAGTTGCTCTGAAAGGGCTTGC	3960
QY	2881	CACGGTAGTCATAATTAGGCCATTGGGACTCGGCCATGAGCTGGCTTTATCC	2940	QY	3961	CNCCTGATATTTTATCATGCTCTCATGCTCTATTTTTAACTCTGGAGCAGTAGTTTATAC	4020
Db	2881	CACGGTAGTCATAATTAGGCCATTGGGACTCGGCCATGAGCTGGCTTTATCC	2940	Db	3961	CNCCTGATATTTTATCATGCTCTCATGCTCTATTTTTAACTCTGGAGCAGTAGTTTATAC	4020
QY	2941	ACAGTGATGGACTCGCCCTCTGGGTGAGACTCTAACAGGAAAGATAGACGTGA	3000	QY	4021	TTATGATTATTAAATTATAGAAATAATTGATTAAATTGATTAAATTGATTAAAGGG	4080
Db	2941	ACAGTGATGGACTCGCCCTCTGGGTGAGACTCTAACAGGAAAGATAGACGTGA	3000	Db	4021	TTATGATTAAATTGATTAAATTGATTAAATTGATTAAATTGATTAAAGGG	4080
QY	3001	GCACATGGTTAAATTGGGAACTTGTGCTGAGACTCTAACAGGAAAGATAGACGTGA	3060	QY	4081	TTAAATTATATATGTTAAATTGAAAGTACATTTTTAAACTTCACTTCTGCTTCAAGCT	4140
Db	3001	GCACATGGTTAAATTGGGAACTTGTGCTGAGACTCTAACAGGAAAGATAGACGTGA	3060	Db	4081	TTAAATTATATGTTAAATTGAAAGTACATTTTTAAACTTCACTTCTGCTTCAAGCT	4140
QY	3061	ACTGGACTTTAGAAACTAAAGCATGCTGCTTATATCATTCTGTGAAATTG	3120	QY	4141	TTTTGATTAGGAGAAATTCTGAAAGTAGGCTTTCTCTGCTTCAATTACTGCTT	4200
Db	3061	ACTGGACTTTAGAAACTAAAGCATGCTGCTTATATCATTCTGTGAAATTG	3120	Db	4141	TTTTGATTAGGAGAAATTCTGAAAGTAGGCTTTCTCTGCTTCAATTACTGCTT	4200
QY	3121	ATATGTTTTCTTCCCTTAAGAATCAAAGAAAAACACAGTCCTACTCCG	3180	QY	4201	CCTTGTGTTCTATGAAACATCAAACTGAACTTCTGAACTTCTGCTTCA	4260
Db	3121	ATATGTTTTCTTCCCTTAAGAATCAAAGAAAAACACAGTCCTACTCCG	3180	Db	4201	CCTTGTGTTCTATGAAACATGCTCTGAACTTCTGCTTCAATTACTGCTT	4260
QY	3181	ATGTTATCTCTGTGTCATGAAACTCTGAACTCTGAACTTCTGAACTTCTGAA	3240	QY	4261	ATATGCTTAAATGAAATCAAATGAACTTCTGAACTTCTGCTTCAAAATAAAC	4320
Db	3181	ATGTTATCTCTGTGTCATGAAACTCTGAACTCTGAACTTCTGAACTTCTGAA	3240	Db	4261	ATATGCTTAAATGAAATCAAATGAACTTCTGAACTTCTGCTTCAAAATAAAC	4320
QY	3241	TTGAAGTTAAATGCAACTTCACTGCACTTCACTGCAACTTCACTGCA	3300	QY	4321	CTATTTTAAATGAAATCAAATGAACTTCTGAACTTCTGCTTCAAAATAAAC	4380
Db	3241	TTGAAGTTAAATGCAACTTCACTGCAACTTCACTGCAACTTCACTGCA	3300	Db	4321	CTATTTTAAATGAAATCAAATGAACTTCTGAACTTCTGCTTCAAAATAAAC	4380
QY	3301	AAATATACAGTAAATTCGCACTTCACTGCACTTCACTGCAACTTCACTGCA	3360	QY	4381	ATGTTACTTAAGAAGTCAATTGAACTTCACTGCACTTCACTGCAACTTCACTGCA	4440
Db	3301	AAATATACAGTAAATTCGCACTTCACTGCAACTTCACTGCAACTTCACTGCA	3360	Db	4381	ATGTTACTTAAGAAGTCAATTGAACTTCACTGCACTTCACTGCAACTTCACTGCA	4440

Db	6601	CATGTGCACITGGAAAAGGAGATCTGGACATGCCAATGTATTACAGAACAAACCCAGCTT	6660	
Qy	6651	TTCCTCAGAAGGTGACTCTAACTCTAGAAATTGGGCTTTATAAGGGTGTGGAGT	6720	
Db	6661	TTCCTCAGAAGGTGACTCTAGAAATTGGGCTTTATAAGGGTGTGGAGT	6720	
Qy	6721	GAAGAACATGCTGTATGTTACTAACAGCCCTTGAAATTACAAACCTGGAAATCCATT	6780	
Db	6721	GAAGAACATGCTGTATGTTACTAACAGCCCTTGAAATTACAAACCTGGAAATCCATT	6780	
Qy	6781	AGGAAACCGATTCATACCTGCTATGTTACTAACAGCCCTTGAAATTACAAACCTGGAAATCCATT	6840	
Db	6781	AGGAAACCGATTCATACCTGCTATGTTACTAACAGCCCTTGAAATTACAAACCTGGAAATCCATT	6840	
Qy	6841	AATGAAAAAGTGTGACTAGTACTCTAAAAATGTCTTAATGATAAAGTTTGAGTCAA	6900	
Db	6841	AATGAAAAAGTGTGACTAGTACTCTAAAAATGTCTTAATGATAAAGTTTGAGTCAA	6900	
Qy	6901	ATAGAAAAGAAAAAAATCTGCATTCCAGCGGAATTGGTATATTATTCATGCATTAA	6960	
Db	6901	ATAGAAAAGAAAAAAATCTGCATTCCAGCGGAATTGGTATATTATTCATGCATTAA	6960	
Qy	6961	AATTGCTPATTCGTAAATGGAAATTCAACTGGCTTATCATGTATATCGGTACTTAA	7020	
Db	6961	AATTGCTPATTCGTAAATGGAAATTCAACTGGCTTATCATGTATATCGGTACTTAA	7020	
Qy	7021	ATGTATTCACAACACTACTGTGATTGTGATAAAATAGACAAAAGTCATATTTTTGT	7080	
Db	7021	ATGTATTCACAACACTACTGTGATTGTGATAAAATAGACAAAAGTCATATTTTTGT	7080	
Qy	7081	GTTGTATAAGCTCTGTAAATAGGAACTACATTAGGAACTGCGACTACATTT	7140	
Db	7081	GTTGTATAAGCTCTGTAAATAGGAACTACATTAGGAACTGCGACTACATTT	7140	
Qy	7141	AAACATTCACTCCAAAGAGCACTATTATGGCATTACCGATAAAATTA	7200	
Db	7141	AAACATTCACTCCAAAGAGCACTATTATGGCATTACCGATAAAATTA	7200	
Qy	7201	ATTTGCTGCTAAATAATAATAGTACTGAGCTTCTGGCCTACAGTGTATGTTGC	7260	
Db	7201	ATTTGCTGCTAAATAATAATAGTACTGAGCTTCTGGCCTACAGTGTATGTTGC	7260	
Qy	7261	TGTAAGGATAAGATACTGTGAAATTCCACAAATAATATGAAATAATCTGCC	7313	
Db	7261	TGTAAGGATAAGATACTGTGAAATTCCACAAATAATGAAATAATCTGCC	7313	
RESULT	3	AC107952/C	142102 bp	DNA linear
LOCUS		Homosapiens chromosome 8, clone RP1-140116,	PRI 23-JUL-2002	complete sequence.
DEFINITION		Homosapiens chromosome 8, clone RP1-140116,		
ACCESSION		AC107952		
VERSION		AC107952.5	GI:21930261	
KEYWORDS		HRCG.		
SOURCE		Homosapiens (human)		
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE		Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,		
AUTHORS		Brown,A., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,		
TITLE		Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,		
JOURNAL		Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,		
REFERENCE		Ferreira,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,		
AUTHORS		Ginde,S., Gord,S., Goettet,M., Graham,L., Grand-Pierre,N.,		
		Hagos,B., Hulten,L., Huime,W., Iliev,I., Johnson,R., Jones,C.,		
		Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R.,		
		Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,		

FEATURES

McDonald, P., Majjor, J., Marquis, N., Matthews, C., McCarthy, M., McEvans, P., McKernan, K., Meldrim, J., Maneus, I., Mihova, T., Mlenga, V., Murphy, T., Nguyen, C., Nicol, R., Norbu, C., Norman, J.J., O'Connor, T., O'Neil, D., Oliver, J., Peterson, K., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talanias, J., Tesfay, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Vieil, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
3 (bases 1 to 142102)	
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Grand-Pierre, N., Hagos, B., Hulten, L., Iliev, I., Jones, C., Kamat, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, D., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, I., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spangler, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfay, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vieil, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
REFERENCE	
AUTHORS	
4 (bases 1 to 142102)	
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Grand-Pierre, N., Hagos, B., Hulten, L., Iliev, I., Jones, C., Kamat, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, D., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, I., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spangler, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfay, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vieil, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
TITLE	
JOURNAL	
4 (bases 1 to 142102)	
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Grand-Pierre, N., Hagos, B., Hulten, L., Iliev, I., Jones, C., Kamat, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, D., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, I., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spangler, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfay, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vieil, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
REFERENCE	
AUTHORS	
5 (bases 1 to 142102)	
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Grand-Pierre, N., Hagos, B., Hulten, L., Iliev, I., Jones, C., Kamat, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, D., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, I., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spangler, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfay, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vieil, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
TITLE	
JOURNAL	
5 (bases 1 to 142102)	
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Grand-Pierre, N., Hagos, B., Hulten, L., Iliev, I., Jones, C., Kamat, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, D., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, I., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spangler, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfay, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vieil, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
COMMENT	
On Jul 23, 2002 this sequence version replaced gi:21699524. All repeats were identified using RepeatMasker:	
Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMaster.html	
----- Genome Center / MIT Center for Genome Research	
Center: Whitehead Institute / MIT Center for Genome Research	
Center Code: WIBR	
Web Site: http://www-seq.wi.mit.edu	
Contact: sequence_submissions@geneome.wi.mit.edu	
----- Project: RepeatMasker	
Center project name: L2490	
Center clone name: 140_I_16	
----- Location/Qualifiers	

source 1. .142102
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="Taxon:9606"
 /chromosome="8"
 /map="g"
 /clone="RP11-14O16"
 /clone.lib="RPCI-11 Human Male BAC"
 /repeat_region complement(416. .691)
 /rpt_family="AluJb"
 repeat_region 1097_1215
 /rpt_family="TCTA.n"
 repeat_region complement(1495. .357)
 /rpt_family="L1PA10"
 repeat_region complement(3983. .4444)
 /rpt_family="MLR2B1"
 complement(4435. .5231)
 /rpt_family="ERVL-B1"
 repeat_region complement(5235. .5457)
 /rpt_family="ERVL"
 repeat_region 5464_7077
 /rpt_family="MERS2A"
 complement(7079. .7553)
 /rpt_family="ERVL"
 repeat_region complement(7554. .7707)
 /rpt_family="L1MC4a"
 complement(7716. .7912)
 /rpt_family="Alusg/x"
 complement(7916. .8284)
 /rpt_family="L1MC4a"
 complement(8362. .8565)
 /rpt_family="HAL1"
 repeat_region 8566_8883
 /rpt_family="Alusc"
 repeat_region 8054_9729
 /rpt_family="L1MC4a"
 complement(9745. .15888)
 /rpt_family="L1PA5"
 repeat_region 15897_16229
 /rpt_family="L1MC4a"
 repeat_region 16229_16797
 /rpt_family="L1MC/D"
 repeat_region 16798_16831
 /rpt_family="CA.n"
 repeat_region 16838_17389
 /rpt_family="L1MC/D"
 repeat_region 17423_17444
 /rpt_family="(A)n"
 repeat_region 17474_18130
 /rpt_family="L1MC/D"
 repeat_region 18138_18182
 /rpt_family="(CA)n"
 repeat_region 18341_18565
 /rpt_family="L1ME"
 repeat_region 18581_18889
 /rpt_family="AlusX"
 repeat_region 18946_19102
 /rpt_family="L1ME"
 repeat_region 19371_19681
 /rpt_family="AlusX"
 repeat_region 21480_21528
 /rpt_family="G-rich"
 repeat_region 22177_22199
 /rpt_family="PAT_rich"
 repeat_region 26803_26828
 /rpt_family="(T)n"
 repeat_region 27273_27305
 /rpt_family="(CA)n"
 repeat_region 27755_27794
 /rpt_family="PAT_rich"
 repeat_region 27902_27929
 /rpt_family="(TG)n"
 repeat_region complement(33140. .33435)

/rpt_family="AluJ0"
 35118_35145
 /rpt_family="PAT_rich"
 complement(38258. .38621)
 repeat_region complement(39628. .39944)
 /rpt_family="AlusX"
 41172_41289
 /rpt_family="L12"
 repeat_region 42616_42641
 /rpt_family="(A)n"
 complement(44033. .44353)
 repeat_region 43354_45125
 /rpt_family="L1PA7"
 repeat_region 44354_45125
 /rpt_family="L1PA7"
 repeat_region 46756_47032
 /rpt_family="L1ME4A"
 repeat_region 47322_48209
 /rpt_family="L1ME1"
 repeat_region 48210_48557
 /rpt_family="AlusX"
 repeat_region 48558_48920
 /rpt_family="L1ME1"
 repeat_region 48848_48886
 /rpt_family="L1MA9"
 repeat_region complement(49958. .50264)
 /rpt_family="AlusX"
 repeat_region complement(50322. .50607)
 /rpt_family="AluS_q"
 repeat_region 50969_51157
 /rpt_family="AlusX"
 repeat_region 51157_51157
 Query Match 89.2%; Score 6520; DB 9; Length 142102;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 6578; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

QY	719 AAAGGCACATGGCTACTCATTCCTCTGAGAAAACCACAGTGAAATTATGTGAGAAAA 778
Db	31099 AAAGGCACATGGCTACTCATTCCTGAGAAAACCACAGTGAAATTATGTGAGAAAA 31040
QY	779 TGTGTTAAGTGCAGAAAGTACATCAGTGAAACTACATAACAGCCATACAGCTTAACGTCACT 838
Db	31039 TGTGTTAAGTGCAGAAAGTACATCAGTGAAACTACATAACAGCCATACAGGACCTAACAGAA 30980
QY	839 CGTTTAAGTGCAGAAAGTACATCAGTGAAACTACATAACAGCCATACAGCTTAACGTCACT 898
Db	30979 CGTTTAAGTGCAGAAAGTACATCAGTGAAACTACATAACAGCCATACAGGACCTAACAGAA 30920
QY	899 TGGCCTTGCATGCCAAAGTGGTGAACCTCACCTGTAAGGTTGACCTCACCTGACCTTGTGAACTTTTG 958
Db	30919 TGGCCTTGCATGCCAAAGTGGTGAACCTCACCTGACCTTGTGAACTTTTGACCTTGTGAACTTTG 30860
QY	959 AAACACCGGAGTCCTCGGAGAACCTAAATCTCAGGAAAGTGTCTGTTGGGG 1018
Db	30859 AAAGACGGAGTCCTCGGAGAACCTAAATCTCAGGAAAGTGTCTGTTGGGG 30100
QY	1019 TTAAAGAAAAAAACACCGTGCACATTTGTGATCGCGGGTCTACACCCGAAGGATG 1078
Db	30799 TTAAAGAAAAAAACACCGTGCACATTTGTGATCGCGGGTCTACACCCGAAGGATG 30740
QY	1079 TCCCGAGAACATGCTGCACTGTGAAAGGAACCTCCCTGTGATGTTGTGCACT 1138
Db	30739 TCCCGAGAACATGCTGCACTGTGAAAGGAACCTCCCTGTGATGTTGTGCACT 30680
QY	1139 AGGATTTGGCGGAGGATCCTGACCTGACATATGAGAACAGTCAATAAAGAGC 1198
Db	30679 AGGATTTGGCGGAGGATCCTGACCTGACATATGAGAACAGTCAATAAAGAGC 30620
QY	1199 TTCTGAAAGTCAAACAGAACCTGGATTCCTGACCTGACTGACTGATGAGAACGTCAATAAAGAGC 1258
Db	30619 TTCTGAAAGTCAAACAGAACCTGGATTCCTGACCTGACTGACTGATGAGAACGTCAATAAAGAGC 30560
QY	1259 TGCTATAAAGAACAGAGCTCCTTCGGTATGCTTACCTTCAGTGAACCTTATCAAA 1318

Db	30559	TGCCTATAAACAGCAGGCTCCCGGTGATGTCAGTGAACGTGTTACAA	30500	Qy	2399	TGTGAATAACTCAAGGGTTAAATTCTTACTAGCCTCTAAATGGATTAAATCACTAAGT	2458
Qy	1319	AGCCATTCAAACAACTTGCAGTTAAACCTCTACACACATCCATTCAGTCAGA	1378	Db	29420	TGTGAATAACTCAAGGGTTAAATTCTTACTAGCCTCTAAATGGATTAAATCACTAAGT	29361
Db	30499	AGCCATTCAAACAACTTGCAGTTAAACCTCTACACACATCCATTCAGTCAGA	30440	Qy	2459	GCTTCAAATGATTAGACTCCAGTCCAGTGGAGATAATAATGTTGTTGATACACAT	2518
Qy	1379	GCTGGGATCTGCCAACAAATGATCACAACTTACCTTGGAAAGCATGCCCATAG	1438	Db	29360	GCTTCAAATGATTAGACTCCAGTGGAGATAATAATGTTGTTGATACACAT	29301
Db	30339	GCTGGGATCTGCCAACAAATGATCACAACTTACCTTGGAAAGCATGCCCATAG	30380	Qy	2519	AATTTCAGATCAGTATATTCTGAAGACTCTGTGTTCTGGCTAAATAATTGCCATCTT	2578
Qy	1439	ATATGGACACTGTTCATCCCTCACCAACCTTCAAAATAATCCCTTCAAGTCTACCT	1498	Db	29300	AATTTCAGATCAGTATATTCTGAAGACTCTGTGTTCTGGCTAAATAATTGCCATCTT	29241
Db	30379	ATATGGACACTGTTCATCCCTCACCAACCTTCAAAATACTCGTTCAAGTCTACCT	30320	Qy	2579	TATTATGAGCCTTAAGGAAAACACCCTAACACAAAGGATCAGTATTTATAGGAAA	2638
Qy	1499	CATATGAAATTCTATTCCTGAAAAAGAACAGCCATTAAAGGGAAAATTGAGAATTAC	1558	Db	29240	TATTATGAGCCTTAAGGAAAACACCCTAACACAAAGGATCAGTATTTATAGGAAA	29181.
Db	30319	CATATGAAATTCTATTCCTGAAAAAGAACAGCCATTAAAGGGAAAATTGAGAATTAC	30260	Qy	2639	AGAGACTCTGTTAGGGACATGGCATGGCATTCTGTGTCACCTTAATAGTGTGCTCTAAATTGATA	2698
Qy	1559	TGATGGAGTTACAAGGTGGCCCTCTCATCCCAAGAATCTCAAGCATCTCATCAT	1618	Db	29180	AGAGACTCTGTTAGGGACATGGCATGGCATTCTGTGTCACCTTAATAGTGTGCTCTAAATTGATA	29121.
Db	30259	TGATGGAGTTACAAGGTGGCCCTCTCATCCCAAGAATCTCAAGCATCTCATCAT	30200	Qy	2699	CACAGGATAATTGGTGTGTTCATCTCTTAACTAGTGTGTTCATCTCTTAACTAGTGTGCTCTAAATTGATA	2758
Qy	1619	CTAAGCTTAGGTTGGATCCTCAGATGGTCCTAATGAGTGCGAGAACCTCTCC	1678	Db	29120	CACAGGATAATTGGTGTGTTCATCTCTTAACTAGTGTGCTCTAAATTGATA	29061.
Db	30199	CTAAGCTTAGGTTGGATCCTCAGATGGTCCTAATGAGTGCGAGAACCTCTCC	30140	Qy	2759	TAATAGTGTATGGCATGGGATCTCTGAGATATAAGTGTGATCTCTTAACTGCTCTT	2818
Qy	1679	TATCCAAAAGCTCTATCATCATCATCATGTCACCCCCCTAAACACACCCCTAA	1738	Db	29060	TAATAGTGTATGGCATGGGTTCTTAACTAGTGTGATCTCTTAACTGCTCTT	29001.
Db	30139	TATCCAAAAGCTCTATCATCATGTCACCCCCCTAAACACAGATGGTATTTC	30080	Qy	2819	TCCACCTTTACATCCAAGGGTGGCTCAAGTGCACCCAGAAATTCTAACGAGGATAGACGTG	2878
Qy	1739	AGTTGTTTAATTCTACACCTTAAATGGTCTCTCATCATGTCACCTGGGAGCC	1798	Db	29000	TCCACCTTTACATCCAAGGGTGGCTCAAGTGCACCCAGAAATTCTAACGAGGATAGACGTG	28941.
Db	30079	AGTTGTTTAATTCTACACCTTAAATGGTCTCTCATCATGTCACCTGGGAGCC	30020	Qy	2879	CCACAGCTTGTGACTCAGCCCTCTGGTGCACAGACTCTAACGAGGATAGACGTG	2938
Qy	1799	TTGGATGAGCTATTCCAGGAAGGCAATTCTCTGCTCCAGCTCCAAACAAA	1858	Db	28940	CCACAGCTTGTGACTCAGCCCTCTGGTGCACAGACTCTAACGAGGATAGACGTG	28881.
Db	30019	TTGGATGAGCTATTCCAGGAAGGCAATTCTCTGCTCCAGTCCCTCCACAA	29960	Qy	2939	CCACAGCTGTTACATTGGGAAACTCTGAGCTTGTAGGGTACAGCTTGTGACTCAGCC	2998
Qy	1859	CAAGAGATCTCAGGATCTCAGGATCTCAGGATCTCAGTGTCACTGTCAG	1918	Db	28880	CCACAGCTGTTACATTGGGAAACTCTGAGCTTGTAGGGTACAGCTTGTGACTCAGCC	28821.
Db	29959	CAAGCTTCAAGGAGCTTAAAGCAAGTACCAACCTTCCACGTTTCCATCAAGCTTTC	29900	Qy	2999	AAGCACATGGTTACATTGGGAACTCTGAGCTTGTAGGGTACATGCCCTGTAGCCAGGTTAA	3058
Qy	1919	CAGCTTCAAGGAGCTTAAAGCAAGTACCAACCTTCCACGTTTCCATCAAGCTTTC	1978	Db	28820	AAGCACATGGTTACATTGGGAACTCTGAGCTTGTAGGGTACATGCCCTGTAGCCAGGTTAA	28761.
Db	29899	CAGCTTCAAGGAGCTTAAAGCAAGTACCAACCTTCCACGTTTCCATCAAGCTTTC	29840	Qy	3059	AAACTGGACTTTTAAAGGTTAAAGTAAACCATGGCTTAACTGGCTTAACTGGTAACTTGTGCAATT	3118
Qy	1979	AGTAGGATTCTGGGACATGGATTCAATTACGAAATTCTGCTTGTGTTAAATAC	2038	Db	28760	AAACTGGACTTTTAAAGGTTAAAGTAAACCATGGCTTAACTGGTAACTTGTGCAATT	28701
Db	29839	AGTAGGATTCTGGGACATGGATTCAATTACGAAATTCTGCTTGTGTTAAATAC	29780	Qy	3119	TGATATGGTTTCTCCCTTAAGATCAAAACAGCTCTACTC	3178
Qy	2039	CCATTTTTATTAGTGTCCPACTTAAAAGCTTAACTTGTGTTAACTTGTGTTAAATAC	2098	Db	28700	TGATATGGTTTCTCCCTTAAGATCAAAACAGCTCTACTC	28641.
Db	29779	CCATTTTTATTAGTGTCCPACTTAAAAGCTTAACTTGTGTTAACTTGTGTTAAATAC	29720	Qy	3179	CGATSTTATCTTCTGATCAATGTGAACTTCTGGTAAATTTGGATGAGA	3238
Qy	2099	AAATTCTCATTAAGCCAGTATAATAAGAAACTAGCTTAAACTTGTGTTGGGAACTAT	2158	Db	28640	CGATSTTATCTTCTGATCAATGTGAACTTCTGGTAAATTTGGATGAGA	28581.
Db	29719	AAATTCTCATTAAGCCAGTATAATAAGAAACTAGCTTAACTTGTGTTGGGAACTAT	29660	Qy	3239	ATTGAGTAACTGATTAGAAGTAAACTCTGTGAACTACCAAGGTTAACTGACT	3298
Qy	2159	TTGTGTTCTAGTTAACTTAACTGGSTATTTGTCTGATCTGCCTGTCACATT	2218	Db	28580	ATTGAGTAACTGATTAGAAGTAAACTCTGTGAACTACCAAGGTTAACTGACT	28521
Db	29659	TTGTGTTCTAGTTAACTTAACTGGSTATTTGTCTGATCTGCCTGTCACITGCCAATGTGACATT	29600	Qy	3299	AGAAATATAACAGCTAAATCCACACTTCTGGAAATGATAGGAACTATTG	3358
Qy	2219	TAAGACTTTTCTCCATATAGAAACCCATTATTAGTAAACTTCAAAATCC	2278	Db	28520	AGAAATATAACAGCTAAATCCACACTTCTGGAAATGATAGGAACTATTG	28461
Db	29599	TAAGACTTTTCTCCATATAGAAACCCATTATTAGTAAACTTCAAAATCC	29541	Qy	3359	AAATAAGTGTGAGTTGTAGGGTAAACAAACTCATCTGGTAACT	3418
Qy	2279	CATTTCAAATTACTTAACTTAACTGGSTATTTGTCTGATCTGCCTA	2338	Db	28460	AAATAAGTGTGAGTTGTAGGGTAAACAAACTCATCTGGTAACT	28401
Db	29540	CATTTCAAATTACTTAACTTAACTGGSTATTTGTCTGATCTGCCTA	29481	Qy	3419	ATGAAATAACAAATTGAGATAATTATGAACTTAACTGAAATTGAGATAATTG	3478
Qy	2339	CCTTTTGACACTGACTCATTAAGGAAATTAGCATAGTAAATTGAGATAATTGAGATAATTG	2398	Db	28400	ATGAAATAACAAATTGAGATAATTGAGATAATTGAGATAATTGAGATAATTG	28341
Db	29480	CCTTTTGACACTGACTCATTAAGGAAATTAGCATAGTAAATTGAGATAATTGAGATAATTG	29421				

Db	26183	TTCAGAAACCTCTTATAGGTATGCCTTTAAAGCTTATAATTCAAGTGC	26124	Qy	6778	ATTAGAAACGGATTGACATCACCTGACATAGCTGACTGAAATTGATTTTA	6837	
Qy	5698	CATTAGATGATATATGTTAGGCTTGTGATATAATGTTGACAAAATGGTAGA	5757	Db	25043	ATTAGAAACGGATTGACATCACCTGACATAGCTGACTGAAATTGATTTTA	24984	
Db	26123	CATTAGATGATATATGTTAGGCTTGTGATATAATGTTGACAAAATGGTAGA	26064	Qy	6838	GCTAATGAAAGTGTGTTCTATCAATTGTCAGTAACTCTTCAATGATAAAGTTGAGT	6897	
Qy	5758	TGGTATTAAACAGGTACATTTCAGTTTCACTGTTCTTCAATTGTCAGTAACTCTTCAATGATAAAGTTGAGT	5817	Db	24983	GCTAATGAAAGTGTGTTCTATCAATTGTCAGTAACTCTTCAATGATAAAGTTGAGT	24924	
Db	26063	TGGTATTAAACAGGTACATTTCAGTTTCACTGTTCTTCAATTGTCAGTAACTCTTCAATGATAAAGTTGAGT	26004	Qy	6898	CAAATGAAAGAAAAAATCTGATCCAGGGCAATTGTTATTTTATTGATT	6957	
Qy	5818	AATCAGTGTTGTTCTTCAATAGGTATTTCACAACTTACAGTTTACAGTTACGT	5877	Db	24923	CAAATGAAAGAAAAAATCTGATCCAGGGCAATTGTTATTTTATTGATT	24864	
Db	26003	AATCAGTGTTGTTCTTCAATAGGTATTTCACAACTTACAGTTTACAGTTACGT	25944	Qy	6958	AAAATGAAAGAAAAAATCTGATCCAGGGCAATTGTTATTTTATTGATT	7017	
Qy	5878	GTCTCAAGCACACTGTTCCCGAGTGAAGTTAACCAAGTCAACAGTTG	5937	Db	24853	AAAATGTTGCTTATCTGTAATATGGAAATCAGTGGCTTATCATGTTATCTGTTACT	24804	
Db	25943	GTCTCAAGCACACTGTTCCCGAGTGAATCAACCAAGTCAACAGTTG	25884	Qy	7018	AAAATGTTGCTTACAAACTACTGTTGATTGTTGAAATGACAAAGATCATTTTT	7077	
Qy	5938	CITCTAGCTACTGAGATCAGTGCAGATCTTAAATGTTAGTATCTG	5997	Db	24803	AAAATGTTGCTTACAAACTACTGTTGATTGTTGAAATGACAAAGATCATTTTT	24744	
Db	25883	CITCTAGCTACTGAGATCAGTGCAGATCTTAAATGTTAGTATCTG	25824	Qy	7078	TGTGTTGCTTAAATGAGCTTGTGAAATGACAAAGTACTATGAGCTGAGTACATAC	7137	
Qy	5998	TCTTGTGAAATGCTGGTTGTAGTTCTGTAATGTTGCTTAAGTCTG	6057	Db	24743	TGTGTTGCTTAAATGAGCTTGTGAAATGACAAAGTACTATGAGCTGAGTACATAC	24684	
Db	25823	TCTTGTGAAATGCTGGTTGTAGTTCTGTAATGTTGCTTAAGTCTG	25764	Qy	7138	TTAACGATTACATCCAAGAGGAGACTATTGTCATATACCAGATTAAATA	7197	
Qy	6058	TTTTCTCAACTCTGTCAGGCCATCTCTAACACAGATAAGCTGTTGTCACCA	6117	Db	24653	TTAAAGATTACATCCAAGAGGAGACTATTGTCATATACCAGATTAAATA	24624	
Db	25763	TTTTCTCAACTCTGTCAGGCCATCTAACACAGATAAGCTGTTGTCACCA	25704	Qy	7198	TAATTTGCTGCTTAATTAATTAATGACTGAGCTTCTTGGCTTACAGTGTGTT	7257	
Qy	6118	TAGATCPACATAAGTGTGAGTTGCTTAAACCCATAGGGACCTCTTGGATC	6177	Db	24623	TAATTTGCTGCTTAATTAATGACTGAGCTTCTTGGCTTACAGTGTGTT	24564	
Db	25703	TAGATCPACATAAGTGTGAGTTGCTTAAACCCATAGGGACCTCTTGGATC	25644	Qy	7258	TGCTGTTAGATAAGTATGTTGAAATTCCACAAAATATGATAAAAT	7305	
Qy	6178	ATAGAAAGAAATCATGCAACCGATTTGGTAAGGCCACACTGAGCTTGGGC	6237	Db	24563	TGCTGTTAGATAAGTATGTTGAAATTCCACAAAATATGATAAAAT	24516	
Db	25643	ATAGAAAGAAATCATGCAACCGATTTGGTAAGGCCACACTGAGCTTGGGC	25584	Qy	6238	CTTCTATGATGTTGTCACCCCTCTGAGGATGTCAGGCCATGTCCTTGTGTTCTAG	6297	
Db	25583	CTTCTATGATGTTGTCACCCCTCTGAGGATGTCAGGCCATGTCCTTGTGTTCTAG	25524	LOCUS	AC097274	182102 bp	DNA linear	HTG 18-JAN-2002
Qy	6298	CCTGAGAAACGAACTTACAGAACTATTGCTTAATGTTAAACACTCTTCAATTG	6357	DEFINITION	Mus musculus clone RP23-59B17, WORKING DRAFT SEQUENCE, 8 unordered pieces.			
Db	25523	CCTGAGAAACGAACTTACAGAACTATTGCTTAATGTTAAACACTCTTCAATTG	25464	VERSION	AC097274_5	GI:18201773		
Qy	6358	GGTACAGATGAAACGAACTATTGCTTAATGTTAAACACTCTTCAATTG	6477	KEYWORDS	HIG, HTGS PHASEI, HTGS DRAFT.			
Db	25463	GGTACAGATGAAACGAACTATTGCTTAATGTTAAACACTCTTCAATTG	25404	ORGANISM	Mus musculus (house mouse)			
Qy	6418	GGATCATGCTGTTAACTGTAATCTGGGTACAAACTCTCTCTG	6477	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Db	25403	GGATCATGCTGTTAACTGTAATCTGGGTACAAACTCTCTCTG	25344	AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
Qy	6478	CTCATTCAGCTGTTAACTGTAATCTGGGTACAAACTCTCTCTG	6537	JOURNAL	1 (bases 1 to 182102)			
Db	25343	CTCATTCAGCTGTTAACTGTAATCTGGGTACAAACTCTCTCTG	25284	COMMENT	McCombie,W.R., Baker,J.P., Balija,V., Dedhia,N.N., de la			
Qy	6538	ATATATTCTTAACTCTGTTAACTTACAACTTTACCCACACATG	6597	TITLE	Bastide,M., Kuit,K., King,L., Kirchoff,K.A., Miller,B.,			
Db	25283	ATATATTCTTAACTCTGTTAACTTACCCACACATG	25224	JOURNAL	Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R., Santos,L.,			
Qy	6598	ATTCAATGTCACCTGCAAAAGGAGATCTGGCTTACAGAACATGTTACCCAG	6657	COMMENT	Shah,R.S., Spiegel,L.A., Palmer,L., Yang,C., and Zutavern,T.			
Db	25223	ATTCAATGTCACCTGCAAAAGGAGATCTGGCTTACAGAACATGTTACCCAG	25164	TITLE	Mouse Genomic Sequence			
Qy	6658	CTTCTGTCACCTGCAAAAGGAGATCTGGCTTACAGAACATGTTACCCAG	6717	JOURNAL	Unpublished			
Db	25163	CTTCTGTCACCTGCAAAAGGAGATCTGGCTTACAGAACATGTTACCCAG	25104	REFERENCE	2 (bases 1 to 182102)			
Qy	6718	AGTGAAGAACATGCTGTTAACTAACAGGCCCTTGAATTAAACACTGAAATCC	6777	AUTHORS	McCombie,W.R.			
Db	25103	AGTGAAGAACATGCTGTTAACTAACAGGCCCTTGAATTAAACACTGAAATCC	25044	JOURNAL	Direct Submission			

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

59171	b	---ATTGGAACATTTCATTTTGTCAAAATTCTAGAAGTGT 59116	Db
2397	y	-TTCTGTAATTAACCTCAGGGTTAA- ATTCTCTTACTAGCTTCAAATGATTAAAT 2453	Qy
59115	b	GTGTTGTGAAATTAACCTCAGGGTTAA- ATTCTCTTACTAGCTTCAAATGATTAAAT 59056	Db
2454	y	CAA -- GTGCTCAATTAAGTAAAGATCCAGTTGGAAATAAATGTTGTAGAT 2511	Qy
59055	b	TAAGTGTGTTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAT 58996	Db
2512	ACACCATAACTTCAGATCAGTATAATTCTGAGACTCTCTGTCTGGCTTAAATATTG 2571	Qy	
58995	b	ATACATCATTAGTGTGGCTAAATATTGTGAGACT-GCTATGTCAGATCAAATATTG 58937	Db
2572	CCATCTTATTATAGGCCTTAAACAANGCATAGTATTAT 2631	Qy	
58936	b	--CTAAAGAACCTAGAGGTAGGGTAGGGTGTGTGTAGT - - - - - 58879	Db
2632	AGCAAAAGAGACTCTGTAGTGACATGGGATTTGTGTCACTPAATAGTGGCCCTAA 2691	Qy	
58920	b	--CTAAAGAACCTAGAGGTAGGGTAGGGTGTGTGTAGT - - - - - 58879	Db
2692	ATTAGTACACAGGATATTTCGTTGTTTACATGCTTATTTTCATTAA 2751	Qy	
58878	b	ACTAGCCTCAGGCTGGTTCGCAATTTCATCTTCATA----- 58834	Db
2752	ATAATAGTAAATAGTGTATGGCTCATGGTCAATATAGGTAGATCTCTT 2811	Qy	
58833	b	--ACAGTGTGTGCACTCCAAAGGGTGGTCAAGTGGCAGCAATTATTTCTAA 58782	Db
2812	AGTCTTTCCACCTTTCACATCCAAAGGGTGGTCAAGTGGCAGCAATTATTTCTAA 2871	Qy	
58781	b	ATGCCCTGTCACTTCATACCCCA--GGATGGCTTCAGATACCCGTTCAATTGGC 58724	Db
2872	TGTRGCGCCACGGTTAGTCATAATC-TAGAGCCATTTCGGAATGGCAGGG 2930	Qy	
58723	b	ATGGTGACCAAGTATTCCAGAGCTTCAGCTTAAAGCTAAATA 58664	Db
2931	TGTTTATCCACAGCTGGTCACTCGCC--TCGTGGTCAAGACTCTAACAGGA 2987	Qy	
58663	b	TGTGAAACCCPACAATGGCTGGTGCCTCAATGGTGGTCAAGCTAACAGGA 58604	Db
2988	AGATAGCTGAAGCACATGGTCAATTGGGACTGGTCACTGGCCCTGTA 3047	Qy	
58603	b	AGATGGATGAAACA--GCCAAACCCAGAACTTGTCCAGAGATT--TCACATGTT 58551	Db
3048	GCCAGGGTTAAAACCTGGACTTTTGTAGAAGTAAAGCTAGCGCTTATATCATT 3107	Qy	
58550	b	GCCAGGTCAAACCTGGACTTTGTAGTGTCAATGGGACTGGCTCTGTA 58495	Db
3108	CTTGTGTAATTGTGATGTTCTTCCCTTAAAGATCAAAGGAGAAACAAACCA 3167	Qy	
58494	b	CTAAGCAGCTTGGCGGTGAGCTTGGC----- 58465	Db
3168	CAGTCCTACTCCGATGTTCAATGTAATTGCAATTCACCAAACT 3227	Qy	
58464	b	-----TCAGTCGTTACCTTCAGTCAATGTTCAATGTTCAATGCTGTGTT 58434	Db
3228	TGGATGGGAAATTGGAAGTAAATGCAATTAGAAACACTACCTGATGAAACTACCAAACT 3287	Qy	
58433	b	TGGATGGGAAATTGCAATTAGAAACCTACCTGATGAAACTACCAAACT 58377	Db
3288	TTTAAGTCACTGAAATATATACAGTAAATCCCACTTCACTGATCTCGGAAATGAT 3347	Qy	
58376	b	TTTACACAAACCAAGATT----GTAAGCTTAAAGTAAACAAACCTA 3407	Db
3348	AGGATATTGAAATAAGTGTAGTTGTAGGGTAAACAAAGTAAACAAACCTA 3407	Qy	
58327	b	AGGAATGTGAAACCAAGTGTAACTAG-TGGAAACAGGAAATAAGCAAAACCC 58669	Db
3408	TCTTGTGTTAACATGAAATAACAAATTGGAATATATTTCACGTAAATATAGCC 3467	Qy	
57257	b	AGTACCAAGGATTAAATGAAATAACAAATTGGAATATATTTCACGTAAATATAGCC 3467	Db

Qy	4521	TATGTGTTATGTGTCTT-----	4539	-----	Qy	5528	GTCAGCAATACACATATTATAGGCCAATACTTTTAAATAAATAAGCTTGTCACACTC	5587
Db	57199	TGTGTATAATAATAATAATAATAACATATATGTGTATGTATGTATAATA	57140	-----	Db	56159	GTCAGCACTACACA-----TATAGGCCAA-----TATAGGCCAA-----	56110
Qy	4540	-----GGTGTGTGTATAATAACACACATACTAAATCTGCCTGCATG	4591	-----	Qy	5588	TATACATACACATATACAGATAACAGATAACAGATAACAGATAACAGAAAC	5647
Db	57139	TGTGTGTATAATGTGTGTGTGTATATAACATAATCTCTAACCT	57080	-----	Db	56109	TATACATACACATACAGATAACAGATAACAGATAACAGATAACAGAAAC	5650
Qy	4592	AAATT-----AAATACATGGAGGACACATTCTCGGGCACCACTGTTAAATTGGAGTC	4647	-----	Qy	5648	TTCCTTATAGGTATGCCCCTTTPA-TTTTAAGACTTATAT-----AATCAAGTGCCTTAA	5702
Db	57079	TTACCCGGACGGAACTCTAGTACACATCTTCCAGG3AACCTGTAAATGTAGCTC	57020	-----	Db	56049	TTCCTCTATGTTATTTAAAGACTTATATATCCTTCAAGTGCCTTAA	55990
Qy	4648	TTAATTCTATG-TGTAACCTCTTGTGTTCCACCCAGA-CCTGAAATAACAC	4704	-----	Qy	5703	GATGATATATATGAGGCTTCATATAATGCTTGTGTCAC-AAAATGCTAGTGG	5760
Db	57019	TTAATTCTGTGTATGACCTTGTGCTGCCCCGAGCTGTATAACAC	56960	-----	Db	55989	GATGATATATGAGGCTTCATATAATGCTTGTGTCACAAAATGCTAGTGA	55930
Qy	4705	TTCAGAGTAAGGGAAATTCAGCTTAATTGTGTT-----TAAATTGACTGTACTG	4755	-----	Qy	5761	TATTTAAACAGGTACATTTCACAGGTTCTTCAATTGCTTATATGACAGAT	5620
Db	56959	TTCAAAGTAAGGGAAATTCAGCTTGTGTTTTGTGTTATGACTACAGT	56900	-----	Db	55929	TATTTAAACAGGTACATTTCACAGGTTCTTCAATTGCTTATATGACAGAT	55870
Qy	4756	GTCACTAACCCTTTTCAGAGATTCTATTAAGGTGAGGCGAGTCGCTTATTTGA	4814	-----	Qy	5821	CAGTGTGTCTTCAAAAGTTTCAAAAGTTTACAGGTTACAGTTAACCTGTC	5880
Db	56899	GTCACTAACCCTTTTCAGAGATTCTATTAAGGTGAGGCGAGTCGCTTATTTGA	56846	-----	Db	55869	CAGTGTGTCTTCAAAAGTTTCAAAAGTTTACAGGTTACAGTTAACCTGTC	55810
Qy	4815	ATTGACAAATGTTCTAACAGGATGTTCTAACAGAGATGTTCTTCCPTAGAAAAG	4874	-----	Qy	5881	TCAAGACACACTGCTCCAGTCATGTTCCAGTCATGTTAACCTGTC	5936
Db	56845	ATTCACAAATTTAACAGGACTAACAGAGATGTTCTAACAGAGATGTTCTCAGTAGAAA	56786	-----	Db	55809	TCAAGACACACTGCTCCAGTCATGTTAACCTGTC	55750
Qy	4875	AT---TGTGTGTCTATGTCAACTAGATGTTAAATAAGTATTGCCATGCTGTT	4931	-----	Qy	5937	GCTCTAGCCTACTGAGTCATGTCAGTTGAGGAGATCTTAAATGTTAGTATTC	5996
Db	56785	TTCTGTCTACTGTGTTTTGATAGATTGATGTTAACCTTCAACCTT	56726	-----	Db	55749	GCTCTAGCCTACTGAGTCATGTCAGTTGAGGAGATCTTCAATGTC	55690
Qy	4932	TTCATTCTCTAGTGGCCAGATCTTATCTGAAATTCTGTGCTCTGCTGTT	4991	-----	Qy	5997	GTCATTAGGAATGCTGGCTGTAGTTCTGTTAAATGTTGCTTAAATGTTGCTT	6056
Db	56725	TTCATTCTC-----CATGGCTCTGGAATTCTGTGATGCTGCTTGAATGATG	56679	-----	Db	55689	ACCTTAATGTAATGTTCTGCAATAGTTGTTCTGAAATGAGTCTTGTGTC	55630
Qy	4992	TAALAAAAAAALAAAAAAALAAAAAAAGGATAAACATTAATAAAAGTAGTTAGAT	5051	-----	Qy	6057	GTTTTTCACACTGTCAGCAGCCATGTCAGCAGCCATGTCAGCAGTCAGTGCAC	6116
Db	56678	TGAAGAAAGGGTTAACATTACTGATACTGTT-----TATGTGGAT	56633	-----	Db	55629	ATTATTATCATGC-TGTAGCAGCCAGTCATGTTAACATGAGCTTGTGATTTG	55575
Qy	5052	TTGGCCCTAGACAGATAATTGACCTTCACTTCAGTTCACTTCAGTACAGT	5111	-----	Qy	6117	ATAGATCATAGTTGCACTGTTGCACTGTTGCAATGTTGCAATGTTGCA	6176
Db	56632	TTGGACCTAGGGCGATGTGACCCATTCACTTCAGTCACTGCAATGCAA-----	56575	-----	Db	55574	-TAGACCTCATATACTTAAATGTTGTTGATGTTGATGTTGATGTTG	55520
Qy	5112	TAGGTCAACAAACCGGAAGTTGAGTCTGGCACTGTAAGCACTGCTTC	5171	-----	Qy	6177	CATGAAAGAAGTATGTCGAAACAGGATTTGTAAGGGACACTCAGATTGCA	6236
Db	56574	ACGGCATGATCTCGGGCACTGCGGACTAGTGGCGAAAGGCGTGGCACTTCAGTC	56515	-----	Db	55519	TATCGAAAGTAATGTCGTTCTGTAAGGAGATCAACATGTTGCAATTG	55476
Qy	5172	CATGGATATACAGTACCTTATAGTTATGTTACATGTTAAACTTCATTTAAA	5231	-----	Qy	6237	CCPTCTCTAGTGTGTCAAACCCCTCTGAGGATGCGAAAGTCCTTGTATCTCA	6296
Db	56514	CATGGATATTACCAATAACCTTCACTGTT-----CATTYAA	56458	-----	Db	55475	CCPTCTCTAGTGTGTCAAACCCCTCTGAGGATGCGAAAGTCCTTGTATCTCA	55117
Qy	5232	TCCTATTACCAAGTCTAGTTCTAAACCTCAATTGTCCTGGGTGATATGCTACTC	5291	-----	Qy	6297	G---CCTAGAAATGACAGAACTTATGCTAATGCTAATGCTAATGCTT	6353
Db	56457	ACCCATTACTTGTCTCACTGCACTTCAATGCAATGCTAACTGCAATGCAA-----	56398	-----	Db	55416	GATUTCTAAAGATGCGAGGAATGTTACTGATGTTAAAGCATTCATTAAGCT	55357
Qy	5292	TGTGNGCAACTTTT-----ATTTCATTGTTCTTAAGGTGTT-----ATTTC	5347	-----	Qy	6354	TATGGATATGCTGCTTAAGTAACTGCTAATGCTGTTCTCAGAACTTCATG	6473
Db	56397	TGTGTCAGCTTCTTAAACCCATACTGATGCTGTT-----ATTTC	56338	-----	Db	55296	TATGGATATGCTGCTTAAGTAACTGCTAATGCTGTTCTCAGAACTTCATG	55239
Qy	5348	ATTGTTGCTTGTGCTTATTTCATTGCTAATGTTGAAATACTGTT-----	5407	-----	Qy	6474	CTGCTCATTCAGCTGCTATTCTGTTCTCAGAACTTCATGCTGCTAATGCT	6533
Db	56337	ATTGTTGCTTGTGCTTATTTC-TAGTCATTGTTGAAATACTGTT-----	56279	-----	Db	55238	CTGCTCATTCAGCTGCTTAAGTAACTGCTAATGCTGTTCTCAGAACTTCATG	55179
Qy	5408	AATTGGAAAGTAGCTAAATGTTAAAGACTACATGCTGCTTCATCA	5467	-----	Qy	6534	GTAATAATATTATTAATCTGCAACTTCATACACCTTACACACCC	6592
Db	56278	AATTGGACAGTGGCTTAAACCTATACGATTCATGCTTCAATGCTGTTCTCAG	56219	-----	Db	55178	GTAATAATATTATTAATCTGCAACTTCATACACCTTACACACCC	55119
Qy	5468	TTGGTTTAAAGCTAAATGTTAAAGACTACATGCTGCTTCATCA	55227	-----	Qy	6593	AATGTTATCATGTCAGTGGCAAAAGGAGATCTGGACATGAAATGTTAC	6652
Db	56218	TTGGTTTAAAGCTAAATGTTAAAGACTACATGCTGCTTCATCA	56160	-----	Qy			

Page 17

Db	55118	AGTGTATTCTGTTGGCCTTCAAAGGAGATCTTGATGTCATACTGACAA	55059
Qy	6653	CCCAAGCTTGTGTCACAAGGTGACTGTAACAGGCCCTTGAAATAGGG	6712
Db	55058	CCCAAGCTTGTGTCACAAGGTGACTGTAACAGGCCCTTGAAATAGGG	55000
Qy	6713	TGTCGAGTGAAGACATGCTGTATGTTAACTAACAAAACACTGGG	6772
Db	54999	TGTCGAGTGAAGACATGCTGTATGTTAACTAACAAAACACTGGG	54940
Qy	6773	AATCCATTAG-GAAACGGATTCATCATCTGAACATAGCTG-----	6815
Db	54939	AATCCATTAG-GAAACGGATTCATCATCTGAACATAGCTG-----	54880
Qy	6816	----- - GACTGTGAAATGTTAGCTAATGAAAGTGT 6853	
Db	54879	TGTCGAGTGAATGTTAGCTAATGAAAGTGT 54820	
Qy	6854	TGGACTAGTACTCTAAATGGTCTAATGATAAAGTTGAGTCACAAATAGAAAGAAA 6913	
Db	54819	TGGACTAGTACTCTAAATGGTCTAATGATAAAGTTGAGTCACAAATAGAAATGAA 54761	
Qy	6914	AAATATCGCATTCAGGGGAATTTTGTTATTTTAAATTGCTATTCCTG 6973	
Db	54760	AAATATCGCATTCAGGGGAATTTTGTTATTTTAAATTGCTATTCCTG 54701	
Qy	6974	TAATATGGAAATCACTGGCTTATCTGTTATCTGTTACTTAATGTTGACTCACAA 7033	
Db	54700	TAATATGGAAAGCAACTGGTTATCTGTTATCTGTTACTTAATGTTGACTCACAA 54641	
Qy	7034	CATCTGTGATTGTTGATAAAATATGAAAGCATATTTTGTGTGTATAAGCT 7093	
Db	54640	CATCTGTGATTGTTGATAAAATATGAAAGCATATTTTGTGTGTATAAGCT 54588	
Qy	7094	CTGTAATATGAAATCACATTATGAAAGCTGCAGTGATACTACATTCACATC 7153	
Db	54587	CTGTAATATGAAATCACATTATGAAAGCTGCAGTGATACTACATTCACATC 54529	
Qy	7154	CAAAGAGGCAACTATTATGTCCATATACCAGATTAATATTGTGTCTAAT 7213	
Db	54528	CAAACAGGCAACTATTATGTCCATATACCAGATTAATATTGTGTCTAAT 54473	
Qy	7214	TAATATAGTACTGCGACTCTTGTGCCCATACTGTTATGTTGCTGTTAGATAAGA 7273	
Db	54472	TAATATAGTACTGCGACTCTTGTGCCCATACTGTTATGTTGCTGTTAGATAAGA 54414	
Qy	7274	TATGTGAAATTCCAAATATATGAAATAAT 7305	
Db	54413	TGGATGAAATTCTGCAAGATACTTGAGTGTGAAT 54382	
RESULT	5		
AL807387/C	AL807387	241148 bp	DNA linear ROD 03-MAY-2003
DEFINITION			Mouse DNA sequence from clone RP23-24J10 on chromosome 4, complete sequence.
ACCESSION	AL807387		
VERSION	1		
HTGS	AL807387-10		GI:30350042
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
KEYWORDS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	Harrison, E.		
AUTHORS			
TITLE	Direct Submission		
JOURNAL	Submitted (03-MAY-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hunquery@sanger.ac.uk		
COMMENT	Clone requests: clone.request@sanger.ac.uk		
	On May 3, 2003 this sequence version replaced gi:26738130.		
	Sequence from the Mouse Genome Sequencing Consortium. Whole genome shotgun may have been used to confirm this sequence.		
	Sequencing data.		

QY	1139	AGAGATTGCGGAAGGGATCACCTGACTCAGCATATAAGAAGAGTCACAAAGC	1198	2218 TTAAGACTTTTTCATAGGAAAGCCATTAGTAAACTTTAAGTAGTAAACTTTAACAAATC	2277
Db	74786	AGGAGTTGGACAAAGATCACCTGACATAGAGTGACAATCAGAGC	74727	73717 CTCAGAGTTAAATTCTGCTTCAATGGCCATTATTCACCAA-TTTCAGAGTC	73663
QY	1199	TTCGTGAAGTCAAACAGAACCGTGGATTCTTGACCCATTACCTGCAATGTTG	1258	2278 CCATTTCAAATTACTTCTTAAATTCTGATTCTTAACATTAACAGTGGGCTCT	2337
Db	74726	TTCGTGAAGTCAAACGGACCATGTTCTGGACCCATTACCTGCAACATGTTG	74667	73662 CTATGTCAAATTCTGTTAGCTTAAATTCTGCTTAATAGCAATGGCC--	73605
QY	1259	TGCTTATAAAAGAGGAGCTTCAGTGACTCTTGACCTTAACTGTTATAAA	1318	2338 ACCTTTGACATGGCTCATTAACAAATTAGCAATGTTAACTCTTGTAAATGAAAG-	2396
Db	74666	TGCTTATAAAAGTGAACTCCCTGGCCTCATGTCACCTTGICA	74607	73604 ---ATTGAGACTTTTTTGCATAATTAGTACAGTAACTCTTGTAAATGAGT	73549
QY	1319	AGCCATTACAAACACTTTCAGTAAACCTTACACCTTACAGTCATGCCA	1378	2397 --TTGTGAATAACTCAAGGGTTAA-ATTTCCTCTAATGGATAATAT	2453
Db	74606	AGCCATTACAAACACTTTCAGTAAACCTTACACCTTACAGTCATGCCA	74547	73548 GTCGTGAAATACTGAAGTGTTCACATTTCATAGCCACTAAATGGATCATAT	73489
QY	1379	GCGGGGACTGCCCATTACCAATGATCACAACTTACCTTGGGAATGACATGCCA	1438	73544 CAA - GTGCTTCAATGTTAGAGTCACTTGGAAAGATAATAATGTTGTTAGAT	2511
Db	74546	GCTCTGGTGTGTCGTOACCAAATGATAAACCTTACCTTGGGAATGACATGCCA	74487	73488 TAATGGTTTTAAATGAGTAACTGAGTTCAGTTGGGAGATAGTGTCTTAAAT	73429
QY	14339	ATATGGCAACTGTCATCCTCTACCAACCTTCTTAAATCCGGTCACT	1498	2512 ACACCATATTTCAGATCAGTATTCAGTAACTCTGAAAGCTCTGTTGTTCTG	25711
Db	74486	ATATGGATGCTGTCACCCCTCATCTCTGTTCAAATGCCCATTCAGT	74427	73428 ATACCATGTTTCAGATCAATATTTGGAGCT-GCTATTGCTGCTTAAATATTT	73370
QY	1499	CATATGCAATTCTTCTTAACTCTGAAAGACGGCAATTAGGGAAATTGAGTAC	1558	2572 CCATCTTTATTAGGCCCTTAAGGAAACRAACCCAAACAAAGCTCAGATT	2631
Db	74426	CATAGCAACCTCTTCTTCTGAAAGACGGCCATTAGGGAAATTGAGTAC	74367	73369 -----AAGTTTCAGTATCTA-----AAGTTTCAGTATCTA-----	73354
QY	1559	TGATGGAGTTACAGGTGGCGTGCCTCTTCAACAGATCCTCAT	1618	2632 AGCCTAAAGAGACTCTGTTAGGTCACATGGCATTCTGCTGTCACCTTAATAGT	2691
Db	74366	TGATGGAGCTACAGTGGTGCACATTCTCATCTCAGTCCAGATC---TCAT	74310	73353 --CTAAAGAAAACCTAGAGGGTACGAGGTAGCAGGTCTGTTGTTA-----	73312
QY	1619	CTAAGCTAGGTTGATGGCTCTTCAACAGATCCTCTCAGTGGGAACTCTCC	1678	2692 ATTAGTACAGGATATTGTCATGGCTGTTCTAACHTCTCATCCTTCATA-----	73267
Db	74309	CTAAGCTAGGTTGAGGCTTCAAGCTGGCTGCGGACCTCTCCC	74250	73311 ACTGGCCCTOAGGTGGTTGCTCCATCATACTCCTCCATA-----	73211
QY	1679	TATCCAAAGGCTPATCCTCATCAGTGAACCCCTAAACACACCCATTTCTC	1738	2752 ATATAGTAACTATGTTGCACTGGCTTCAAGTGTGATAATAGGTTGCTTT	2811
Db	74249	TGTAAGAAGGCTPATCCTCATCAGTGAACCCCTAAACACACCCATTTCTC	74190	73266 -----AACAGTGTGGCATGAAGTCATCGGGATCTGAGAAGGTTGTGAT	73215
QY	1739	AGTGTGTTAACTTCTATCCCTTAATGGTCTCCCTATACTCCCTATCAGTGGGAGCC	1798	2872 TTGTGGCCACGGTTAGTCATTCATGCAATTGCTGCTGCTGCTGCTGCTG	2871
Db	74189	AGTGTGTTAACTTCTATCCCTTAATGGTCTCCCTATACTCCCTATCAGTGGGAGCC	74130	73214 ATGGCTGTCACCTTCATACCCCA--GGATGGCTCTCATGCTGCTGCTG	73157
QY	1799	TTGAAATGAGCTTCCAGGAAGAGCACATTCTCTGTTCCAGGCTCCACCAA	1858	2873 TAGGTATCCCACTGGTGTGACTGACTGCC--TCTGTGGGTGACTGCTGCTG	2897
Db	74129	TTGGGATGAGCTTCCAGGAAGAGCACATTCTCTGTTCCAGGCTCCACCAA	74070	73096 TGTAACCTCATAGTGGTCCCTTAACTGATGGTCCCTTAACTGAGCTCAGGAA	73037
QY	1859	CACGGATCCTCAGGATCTGCAAACCATATAGGGCTGGGCTCTGCACTCATGTCAG	1918	2988 AGATAGACCTGGAACGACATCTGGTACATGGTAACTGCTGCTGCTGCTG	3047
Db	74069	CACGGATCCTCAGGATCTGCAAACCATATAGGGCTGGGCTCTGCACTCATGTCAG	74010	73036 AGATGGATGAAATA---GCAAAACCCAGAACTTGTGTCAGGAATT	72984
QY	1919	CAGTTTACAGGAGTTTAAGCAAGGTACCCCTCCACAGTCAAGTTTC	1978	3048 CCTGGCTTAAACTGACTGTTGTTCTCTTAAAGCATAGGCTTATATCATT	3107
Db	74009	CAGCCTTACAGGAGCTCAGCTAACGACTACCTTGGCTTCAACAGGGTTC	73950	72927 CTAAGCAGCTCTGGCTGGCTGAGCTCTGCT	72898
QY	1979	AGTAGGATTCTGGGACATGGATTCTTACATAGAAATGATGTTGAGTAA	2038	72983 GCAGAGCTAAACTGAACTCTGGGSCAGGAAATTGCAATTGCTGCTGCTG	3227
Db	73949	AGTAGGATTCTGGGACATGGATTCTTACATAGAAATGATGTTGAGTAA	73992	72897 -----TCAGTGTGTCACCTCTGGGCTGAGTAAAGGCTA---TCACATCAGT	72867
QY	2039	CCATTTTATTCTACTGCTAC-TTAAACAGTATAAAATTCTGTTTGTATA	2097	3228 TTGGATGGAAATTGAACTTAACTGCTTCTGGCTGAGCTTCTGCTGCTG	3287
Db	73891	CCATTATTATCTACTGCTACTTAACTAATCAAACACTCTGCTTGTATA	73932	72866 TTGGATGGAAACTTAACTGAAAGCTTCACTGCTGCTGCTGCTGCTG	32810
QY	2098	CAAATTCTCATTAAGGCACTTAAATAGAAACTGACTCTTAACTGAGCTT	2157	73831 TCAGGTTTATGAGCAGTAAATATAGGAACTAGCTCTCTGATAAGCTT	73717
Db	73771	TCAGGTTTATGAGCAGTAAATATAGGAACTAGCTCTCTGATAAGCTT	73718	3288 TTTAAGTGTGACTGAGTAAATATACTAGTAAATACAGTAAATACAGTAA	3347

REFERENCE	AUTHORS
1 (bases 1 to 278377)	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbera, J., Benton, J., Blamey, K., Blankenburg, K., Bonnin, D.P., Bouck, J., Bowie, S., Brieva, M., Brown, E., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cazavos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davis, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., DeGado, A., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Dubin, K.J., Barnhart, C., Edjaar, D., Edwards, C.C., Elhai, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Gara, N.N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris-K., Hart, M., Haylat, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Hollaway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louiselle, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, K., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwoke, N., Okwonou, G., Oragny, N., Oriedo, R., Pace, A., Payton, B., Perry, J., Perez, L., Peters, L., Pickens, R., Primus, E., Po, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshani, N., Siisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, K., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tameraisa, A., Tang, H., Tansey, J., Taylor, C., Taylor, T., Teirfford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villanova, D., Vinson, R., Wang, Q., Wang, S., Ward, Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wieczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 278377)		
AUTHORS	Worley, K. C.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
REFERENCE	3 (bases 1 to 278377)		
AUTHORS	Worley, K. C.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
COMMENT	On Sep 19, 2002 this sequence version replaced 91:22538586.		
COMMENT	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a "contig-scaffold"). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.		
-----	Genome Center	Baylor College of Medicine	
Center code:	BCM		
Web site:	http://www.hgsc.bcm.edu/		

QY	3234	GGAGATTGAACTTAAATGCTTAACTACCGTGTAACTACCAGACTTAACTACCAAAAGTTTAAG	3293
Db	146507	GGAACCTAAATTAACATCAMAAGAACACTGAGTT--CCAGAGTTACA	146563
QY	3294	TGACTGAAATATACTACGTAATACTCCACTTTCATGCCATCTCTGCAAATGATAGGT	3353
Db	146564	CCACCAAAAGTTG-----TTAAGAGCCACCTTTACACATCCACAGAGCAGGAAAT	146618
QY	3354	ATTCGAAATAAAGTCAGTTGAGGGTACACAGTAAGTAACAAACCTATCTGG	3413
Db	146619	GTTAGAAATAAGTCAAATTATAGGGAAACAGGAAATAACGAAACATGAA	146678
QY	3414	TTACATGAAATAACAAATGAACTATTATAGGCTTTCC	3473
Db	146679	TACGAATGACCAAAAGGATGTGGTCACTGA-----GCAAGTAGCAGGAACT	146731
QY	3474	TCAGATTGACAAACAACTAACTCTCTTAAGGCTAAATTAAATAATTTCTAAGGTT	3533
Db	146732	TCTTGTCAGACAGGCCACCAAGAAATGTCATGGTAGCT-CCGGTAACTGGGT	146790
QY	3534	CAGTGGGACATTAACCTAACATACTTAAGGACTCTGGTAACACCTATTAG	3593
Db	146731	TATTGTCGECATACCTACCTACGAGCTAC-----TAAGGACTCTGG	146946
QY	3594	GACCTTTCGTTGGGAAAGGGCAAGGGTGGGAGGTTTAAGAGATAATCTCT	3653
Db	146847	ACCCCTTCCTGTTGGCA-----GGAGTTTATAGAGGATATPATCTCT	146893
QY	3654	TTAAAAAAAACAGAAAGAAA-----TATTCTGAACTCATTAGGCCT-A	3701
Db	146894	TTAAAGAAAAAGAGAAACTACTTTTATCTGAGACTCTGGCTGC	146953
QY	3702	TATGGAAACTCTCCCTGGCAGATTGCAATCTGCAGATTGCAATGTTTACCA	3761
Db	146554	TCGGGAAGCTCTCCCTGTAAGGGCAGATTGCAATTGCAATTGTTTACO-	147012
QY	3762	AGMATTCTAAATGAGTGTGCAATTACTGAAATAACATATTAAATTTGGGA	3821
Db	147013	-CATCACTCTAAACAAAGATTACTGACATGATGCAATTGCAATTGCGGA	147070
QY	3822	GTTGATATAATTGTTGAAATGTTGAAATAATGTTAAATGGGGCTTCATAATAT	3981
Db	147071	GTTGATGTTAAACTGTCCT-----GTGATGGGGTTCACTATPAT	147123
QY	3882	ATATATAATACACACACACACACATGCAACATACGCAACATGCAATTGCA	3941
Db	147124	GT-----CATACACGGCTAGCTGCACTCATGCAATTGCAATTGAGGTG	147163
QY	3942	CTCTCTGAAAGGGCCTTGGCTCTGATATTTCATGCTCTATTTAACTCTT	4001
Db	147164	CTCCCTAGAGG-GCTGCAAGCTCTGGTCT-----	147198
QY	4002	GGAGGNGTAGTPTTAACATTTAGTTAATGAAAATTACATTTAATGAA	4061
Db	147199	AGAGCAGTAGTTTATACGTTAGTTAATTTTGTGAAAAA-TACATTTAAAT	147257
QY	4062	AAAAGTCGTTCAAGGCTTAAGGCAATTATGTTAATAGGAAGTACATTAAAT	4121
Db	147258	AAGAAAGTGTTCAAAAGCATTAAATTATACGTTAATAGGAAATACTTTT	147315
QY	4182	CTCTGCTTCATTCTGTTCCATTGCTTAACGATTTGCTTAATCTT-T	4240
Db	147376	CATGACGCC-----TTCTTAGTTCTGAAACCGATACTGTTACCTATG	147426
QY	4241	AGTGTGATGTTAAGTGTGTTCAATTGCTTCAAGTGTGCTTCAAGCTGCT	147426
Db	147427	AGTGTGATGTTAAGTGTGTTCAATTGCTTCAAGTGTGCTTCAAGCTGCT	147486
QY	4297	GTTGACAGAGCACAA-----AGAAAACCTTATTAAATGAAATCACAAATGAAAG	4351
Db	147447	GTTGACAGAGCACAA-----AGAAAACCTTATTAAATGAAATCACAAATGAAAC	147546
QY	4352	TGTTGAAACACPACTTTCTAGCTTGTAAATGTTAAATGACTTAAAGTCATCAAATTAGTGAAC	4411
Db	147547	TATGAAACCCACTAGTTTATCTCACTAAATGAACTGACTTCAGAGCACCA	147603
QY	4412	TGAGACATTGGCCTTAGGGCTTAATTTAGGAAATTTGAAAGTGTGCTTCAAGCTGCT	4466
Db	147604	TGAGACGTGGCCTTGTGAGCTGTATATTCAGTAAATGAAAAAGGATGAC	147663
QY	4467	AGGATTTAACTAAACGCTTGGGAATGGCCATATGTTGTTGTTGTTG	4526
Db	147654	AGGATATAAGATAAAACCATTCGCA-----TGCTGTTGTTGTTG	147721
QY	4527	TATGTCGTCGTGTATACTACACACACATACTAAATTTGCTTAAAT	147781
Db	147722	TATATATATATATATGTTATATAATAATAATAATAATAATAATAATA	147781
QY	4586	TGCAATTAACTAACATGAGGGCACATCCTGCAACCTTCACTCTTCAAC	4645
Db	147782	TCTTGCCAGGCAGACTCTAAACACATCTTCCAGGTACCATGTTAAACGTA	147841
QY	4646	TCTTAATT---TTCATCTGTAACCTCTGGCTGTTCCACCCCCAGACT-TGAAATAA	4701
Db	147842	TCTTACTTTCTTTGTTGATGTAACCTCTTGGTGTCTGCCCACAGACTGAA	147901
QY	4702	CACTTCAGAGTAAGGGAAATTCGTTTAAATTGACTGTACTGTCACTT	4760
Db	147902	CATTGAACTAACAGGAAATAAGCAATTCTTGGCTTATTGACTGACTGGTAC	147961
QY	4761	TAACCCCTTTGAGCAATTCTTAAAGATAAGTGGGAGCTCGCTTATGTTCA	4820
Db	147962	TAAACCCCTTTGATGAGTTTATATTTAAATTCAGTCACTTTCATGCA	148021
QY	4821	CAATGGTCTAACAGGAGTGAACAGAAATTGGCTTCTTCCTCTCTCT	4877
Db	148022	CAATGTTTACGGGAGCAACGAACTTGGCAACCTTACCCATGGAAAATCTGT	148021
QY	4878	GTTGTTTCTATGTCACATGATGTTAAATAATAGTATGTTCACTT	4937
Db	148082	CACTCTCTCTCTGAGGTAGCTGTTGCAATTGCTTCTCTCTCTCTCT	148141
QY	4938	CTCTAGTGGCAGAACTATTCCTGAAATTCTGTTGCTTGTCTGTT	4997
Db	148142	CTC-----CATGGTCTCTGGATTCGTTGCAATTGCTTCTCTCTCT	148179
QY	4998	AAAAAAAACAAAAAAAGGGATTACATTAATAAA-----AGTAGTTGAAATTG	5054
Db	148180	-----ANAGAGAAAGGGTTAAATTAACAAACGTTGAGTTGTTGTT	148226
QY	5055	GGCTTCAGAACAGAATTGGACCTTCACTTCAGTTCACTTCACATGTT	5114
Db	148227	GACCTGAGAGTGTGCTGCTGCCCTCTGCTTCACTGCTGCTGCTGCT	148282
QY	5115	GTCACCAAAACAGGAGTGTGCAATTGCACTTCATGCTATCT	5174
Db	148283	GCTTCAGGGTGTGCGCCCTTTGAGAAAGGTCTGGCACTGTGCAATTGTC	148342
QY	5175	TGATGTTAAAGTGTGTTTAACTGTTATAGTTGTTCACTTGTCAAGTTG	5232
Db	148343	TGATATTTCATTAATACCCCTGCAATTGCTTAACTGCTTAAATGCT	148402
QY	5233	CCTTAACTACCAAGTCAGTTTAACTGTTCACTTGTGCTGGCTGATTA	5292
Db	148403	CCCTTA-CCTGTTGAGGTCTTAAATGAACTGCTTCACTGCTTCA	148461
QY	5293	GTTGCAACTTTT-----ATTGTTAGTGTCTTCAAGCTGTGTT	5343
Db	148462	GTTGCACTTTCCTTCTGCTTCAATTGCTTCACTGCTTCACTGCTT	148521
QY	5344	GCCTATTGTTGCTGCTTCTTGTGCTTCAATTGTTAGTGTGTT	5403

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;	AUTHORS
Mammalia; Eutheria; Rodentia; Muridae; Murinae;	
Rattus	
1 (bases 1 to 244955)	
Muzney,D.,Marie., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alisbrooks,S., Amin,A., Anguiano,D., Anyaibechi,V., Aoyagi,A., Ayodeji,I.M., Baca,E., Baden,H., Baldwin,D., Bandaraanake,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenship,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardozo,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacon,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyte,M., Cres,A., De Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Andra,C., Dederich,D., Delgado,O., Denson,S., Deramo,C.C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escott,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorios,E., Gear,K., Gill,R., Grady,M., Guerra,W., Gunatilleke,P., Haaland,W., Hamil,C., Hamilton,C., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,V., Hines,S., Holland,S.L., Hodgson,A., Hopkins,M., Hollands,B., Howell,S.S., Hulyk,S., Hume,J., Idelbierd,D., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpwitz,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levant,J., Lewis,L., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loresuahewa,L., Loulseged,H., Lozaq,R.J., Lu,X., Ma,J., Maheeshwari,M., Mahindarage,M., Mahmoud,M., Mallojy,R., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeilly,T.Z., Meenen,B., Mileavil-Jevic,A., Miner,G., Minja,B., Montemeyer,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakelueh,O., Okononu,G., Olaruppongagorn,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Pondexter,A., Popovic,D., Primus,E., Pu,L.-L., Puaio,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rotas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scheer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sibson,I., Sitter,C.D., Smaja,D., Snead,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemani,K., Vala,R., Vera,V., Villasana,D., Waldron,B., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Wilson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Smith,H.O., Weinstock,G., and Gibbs,R.A.	
TITLE	JOURNAL
Unpublished	JOURNAL
2 (bases 1 to 244955)	JOURNAL
Worley,K.C	AUTHORS
Direct Submission	TITLE
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL
3 (bases 1 to 244955)	JOURNAL
Rat Genome Sequencing Consortium.	AUTHORS
Direct Submission	TITLE
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL
On Sep 18, 2002 this sequence version replaced 91:21671673.	COMMENT
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.edu/projects/rat/). Each contig described in the feature table below represents a scaffold.	COMMENT
assembly ('contig-scaffold'). Within each contig-scaffold,	

QY	3231	GATGGAGAATTGAAAGTTAAATGCAATTAGAAACTACCTGTATGAACCTACACAAAGTTT	3290		Db	220407	TTCGGTGAGAGGACAAGTAATAAAAACCTATTATGAGGTTCAGAACACC	220348
Db	221387	ATGGAGAACCTTAAATATTTAAACATCAGAAACCTGAGCTT	221331		QY	4349	AAGTGTGGAGGCACTACTTTATCTGTTAAATGACTCTTAAGAGTCATCAAATTAGTS	4408
QY	3291	AAGTCACTGAAATAATACGTTAAATCCACPTTCAGTCAGCATCTGGAAATGTATAGC	3350		Db	220347	ACATGAGAACGACTATCTAAATTTCTAAATTAACATTAAGAG--TCACATTATG	220291
Db	221330	ACACCAACCCAAAGTGTAA---AGGCCCACTTTACACATCCAGAGAACGGAG	221276		QY	4409	AACTGAGACATTGGCCTTAGTGGCTGA-----TTCACACTGCTAAAGGGACT	4463
QY	3351	AGTATTGCAAAACTGAGCTTGTAGGGTAAACAAAGTAAACTTACCTCT	3410		Db	220290	ACCTGAGAGCTGGCTCGTAGGTGATAATTACTGCTAAAMAGGGGT	220231
Db	221275	AATGTTAGGAAATAATGTTGAAATTAGAGGAAATAAGGAAACCCAAAT	221216		QY	4464	ACCRGAGATTATTAAGAACGATTGGAATGGGATAACGCACATAATGTATG	4523
QY	3411	TGGTTAACCTGAAATAACATTGAGAAATATTTACTGAAATAATTAGGCTT	3470		Db	220230	ACCAAGATPATATGAAATAAACGATTGCAA-TGGGATGGCCATATATGTTG	220173
Db	221215	GAATPACGATGCAACAGATGCTCACAGA-----SCAAGGTAGAACTGCA	221163		QY	4524	GTGATGTGTTGTTG-GTGTGTGATAATACACACACATACATTAATCTTG	4582
QY	3471	TCCCTCACATAGAACCCACATACTCTTAAGGTTAAATTATTTCTTAAG	3530		Db	220172	GTGTATATATATATATATATATATATATATAATTTATCTCCT	220113
Db	221162	TCCCTCTGTAGACAGGCCCAAGAATGTCATCGGTAGCT-CGGGTTATCTCCTAG	221104		QY	4583	CCCTGCGATGAAATTCAAACTAAACATGAGGACACATCTCAGGGCAC	220053
QY	3531	GGTAGTGGAGATTAACCTAAAGAACATATCTTAAAGCCTTAAACCTTAAU	3590		Db	220112	AATCTTGCCTCAGCAGAACTTAACATCAGCTTCAACATCTCAGGTGTTAAATTG	4642
Db	221103	GGTTATTTGTCGATCAGCTAACAGTAACTGAACTTAAACCTTAC	221048		QY	4643	GAGTCCTTAATTTCAT---GTTGACACTCTTGCCTGTTCCACCCAGA-CITGAA	4698
QY	3591	TAGGACCCCTTCGGTGGGGATGGGGAGTTTACAGAGTATAPATC	3650		Db	220052	GAGTCCTTAATTTCCTGTTGACTACCTCTGTTGACTGAGTCAAGT	219993
Db	221047	TGGACCCCTTCGTTGGGCA-----GGGTTTTATAGAGATATATA	221001		QY	4639	TAACACTTCAGAGTAAGGGGATTAGCTTAATTGTTTAAA-ATTGACTGTGTTG	4757
QY	3651	TCTTTAAACAAACAGAAGAAA-----TATTCTGAGACCTCATTAGCCCT	3700		Db	211992	TGACATTTCAGTAAGGGAAATAAGCAAACTCTTCGCTTTAGCTGAGT	219933
Db	221000	TCTTTAAACAAACAGAAGAAAACATTTTAAATCTGAGATATATA	220941		QY	4758	CACTAAACCCCTTTGAGAGAATTCTTATTAAAGATAGGGAGACTGCTTATTGATT	4817
QY	3701	--ATATGGAAACUTCTCTTCCTTGTAGGGCGAGTTACCTGAGATGCAATTGTTA	3758		Db	211992	CACTAAACCCCTTTGAGAGAATTCTTATTAAAGATAGGGAGACTGCTTATTGATT	219873
Db	220940	GGCTGGGAAAGCTCTCTCTGTAAGGGCAGTTACCTGATATTGTTA	220881		QY	4818	GCAACAGTTCTAACAGGATGTAACAGAATTGCTTTTCCCCTAAGAACATAC	4874
QY	3759	CCGAAATTCTCATAAAATGAGTGCAGATACTGAAATAATCATTAAATATTG	3818		Db	2119872	GCAACAGTTTAACAGAACGCAACAGAACGTTAACATTAACGAAATTGCA	219813
Db	220880	CCCATCACTTAAGCATAAA---CAGATCTGACTGATGAACTTACATGCTG	220824		QY	4875	ATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA	4934
QY	3819	GGAGTAGTATAATTGTTGAGAAATGTAATTGTAATGTAATGAACTCTA	3878		Db	2119812	TGTCACTGCTGTTGTTGTTGAACTGTTGCTTGTAGCATGGCTTGTAGCTC	219753
Db	220823	GGAGTAGTCAA---ATTGAGAGTGTAACTGTCCT--GTGAATGGCCCTTACTA	220771		QY	4935	ATTCCTGTCGCCCCAACTTCTGAAATTCTGAAATTCTGTTGAGTGTGTTG	4994
QY	3879	TATATATAATAACACACACACACACACACACACACACACACACACAC	3938		Db	2119752	ATTCCTGTCGCCCCAACTTCTGAAATTCTGAAATTCTGTTGAGTGTGTTG	219712
Db	220770	TATGT-----CATACACACACACACACACACACACACACACACACAC	220731		QY	4995	AAAAAAAACAAAAAAACGGATAACATTAATAAA---AGTAGTTGAAAT	5051
QY	3939	TTGCTCTCTGAAAGGGAGCTTGGCTCTGATATTATCATGTCCTPATTTTTAATC	3998		Db	2119711	-----AAAAAAAACGGTTAACATTAACAAAC2AGTGTAGTGTGAGT	219668
Db	220730	CTGTCCTCTGAGG-GTCAGTCAGCTGCTGCTGCTGCTGCTGCTGCTG	220696		QY	5052	TGGGCTCAGACAAGATAATTGAAACCTCATTGTTCTGAACTTCTGTTAA	5111
QY	3999	CTGGAGCAAGTAGTTTATACCTTGTAAATTATCATTTAATGAAATTACATT	4058		Db	2119667	TGGACCTCTGAGAACATGCTGGCACTGCTGTTGCTTAACTGCTC	219552
Db	220695	CCTAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	220637		QY	5172	CATTGTTGTTACAGTACCTTTATGTTATGTTATGTTATGTTAACTGTTAA	5229
QY	4059	AAAAAAGTGTCTCCAAAGGGATAAAATTATATGTTATAAGGAATGACATT	4118		Db	2119551	CATTGTTATTTACCACTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2119433
Db	220636	TAAAGAAAGTGTCTCCAAAGGGATAAAATTATATGTTATAAGGAATGACATT	220579		QY	5290	TCTGTTGTCAGGAGTGTCTCCCTTCAAGAAAGCTGCTGTTGTTGTTG	5340
QY	4119	AATTTCATCAACTGCTCTAGCTTGTAGTTGTTGTTGTTGTTGTTGTTG	4178		Db	2119491	AATCCGTTTACTT-GTGTGAGTTTCTGTTAACTGTTAACTGTTAA	211933
Db	220578	AACTCTCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	220468		QY	5341	TTTGGCTTATTGTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	5400
QY	4239	T-TAGTTGAAATGATGTTGCTTGTCAATTGCTTAACTCACA-----	4293		Db	2119432	TCTGTTGTCAGGAGCTTCTTCAAGCTTGTGCTTGTGCTTGTGCTTGTG	2119373
Db	220467	TCTAGTTGTTGATGTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	220408		QY	4294	TTGGGTGACAGAGCACAAAT----AGAAAACCTTATTTAAATGAAATCACAAATAGC	4348

Db	219372	TTTGCCTATGTCGCTGGCCTTATTTTC-TAGTCAATTGGAAATACTGTATA	219314	QY	6465	CAACCTCTCTGTCATTCACTGAAAGACTACATTAG 6524
QY	5441	TCTGTTAATTGGCAAGTAGCCATTAAACCATTAATCTGACTGAAACATGCCA	5460	Db	218261	CACAT - TCTGTCATCATTCTGTTAACCTAGAAAGGACT-CATATGG 218205
Db	219313	TTGTTGAATTGGCACTGACGTTTGTGAACCATATACTGACTGAAACATGCCA	219254	QY	6525	TCTAAAAGTGACATAATATTTATCTGACTTAACTAACAAACTTTCAC 6584
QY	5461	GAGCCGATATGCTTATTAAGCTTAATAGTAAAGGATACATTTTCAAGGATGT	5520	Db	218204	TGTAAGGTGACATAATCTGACTTAACTAACAACTTAC 218145
Db	219253	GAGGTGTTATTAAGCTTAATAGTAAAGGATACATTTTCAAGGATGT	219195	QY	6585	ACCAACAAATGTTCTGACTTCTGGACATGCAAATGTTTAC 6644
QY	5521	TCTATCTAATGCAAAATACACATTTAGCCAAATTTTAAAGGATCTGCTGT	5580	Db	218144	ACCAACAGTGTACTCTGTCGCTGCAATGTTTAC 218085
Db	219194	TCACTCTAATGCAAAATACACATTTAGCCAAATTTTAAAGGATCTGCTGT	219142	QY	6645	AGACAAACCCAGGTTTGCCACAAAGGTGAAAGTGGCTTA 6704
QY	5581	CAACTCTATACACATACATTTAGCAAAATGAACTCAAACTTAAACCTTA	5640	Db	218084	AGACAAACCCAGGTTTGCCACAAAGGTGAAAGTGGCTTA 218025
Db	219141	CAACCTCTATACACATACATTTAGCAAAATGAACTCAAACTTAAACCTTA	219082	QY	6705	TAATAGGCTGTGGGTGAAAGAACCTGGCTATCTACTAACGCCCTTGAAATTAAACA 6764
QY	5641	CAGAAACTTCTCTATGGTTATGCTTA-TTTAAAGCTTAT---AATTCACTG	5695	Db	218024	-CATAGGGTGGGTTGAGAACTGTATAAGTTACTAACGCCCTTGAAATTAAACA 217966
Db	219081	CAGAACTTCTCTATGGTTATGCTTA-TTTAAAGCTTATTTAACCTTCACTG	219022	QY	6765	AAACTGGAAATCCTATTAG GAAGCGATGCAATACCTGCAACATACTGCTGCTGCTG 6823
QY	5696	GCCATTAGTGTATATATGGCCTTGTGATATAATGCTTCTGAAAA-AATGGT	5754	Db	217965	AAACTGGGAATCCTATTAGAAAAATGGATGCAATGCTGAACTAAGCTGTA 217906
Db	219021	GCCATTAGGATATATGGCCTTGTGATATAATGCTTCTGAAACAAATGGT	218962	QY	6824	AAATTGTATTTTGTGACTTAATGAAAGTGTGACTTCAAAATGTTTA 6880
QY	5755	AGAGGTTATTTAACAGGTACATTTCACAGTTCTTCAATTTGCTATATGCA	5814	Db	217905	AAATTTTTTGTGACTTAATGAAAGCTTGGACTGTGACTCTAAAAATGTTTA 217846
Db	218961	AGATGTATTTAACAGGTACATTTCACAGTTCTATGATGTTCTGATATGCA	218902	QY	6881	ATGATAAAGTTGGTGGACTCAAATGAAAGAAAAAATGCAATCCAGCGGAATTG 6940
QY	5815	CAGATCTGTTGCTTCAATGGTTTACATGGTTTACAGGTTTACAGGTTA	5874	Db	217845	ATGTAAGTTGGTGGACTCAAATGAAAGAAAAAATCTGCACTTCCAGACTTAATTG 217786
Db	218901	CAGATCTGTTGCTTCAATGGTTTACAGGTTTACATGGTTTACAGGTTA	218842	QY	6941	TATATTTTTATTGGATTAAAATGGAAATCTGTTAATTGGAAATCTGTTATTCAGTGGTTATC 7000
QY	5875	CGTGTCCTAAAGCACACTGCTTCCAG ---TACGTAACTTAAAGCTTCA	5930	Db	217785	TATATTTTTATTGGCTTTAAATTGCTATCTGTTATTGGATAAAATATAG 217726
Db	218841	CGTAATCTAAAGCACACTGCTTCCAGTCATAGTAAAGTACACTTCA	218782	QY	7001	ATGATATCTGTAATTGATTGATTCAAAACTACTGTTGTTAAATGTTG 7060
QY	5931	CAAGTTCTCTAGCTTCTAGATCTGAACTGCTTCACTGTTCACTGTTCA	5990	Db	217725	ATGATATCTGTAATTGATTCAAAACTACTGTTGTTATTGGATAAAATATAG 217666
Db	218781	CAAGTTCTCTAGCTTCTAGCTACTGCACTGTTCACTGTTCACTGTTCA	218722	QY	7061	ACAAGATCATATTTTTGTGTTGTTGATAAGCTGTAAATAAGAAATACATTATGAG 7120
QY	5991	GTATTGTCATTAGAAATGGCTGGTTAGTTCTGGTAATGTCTGCTTAAGT	-TGA	Db	217665	ACAAGAA-----TGGCTGTGTGTGTTGTCAGTAGCAGTCAGTAGCAG 217614
Db	218721	ATATTCGCCCTCATGAAAGCAGGTGTTTCTCGACAAATGAACTGCTG	218662	QY	7121	CTGGTAGTGTACTCATTTAAACATTCACTCCAAAGGCAACTATTATGTTGTTG 7240
QY	6048	ATTTGCTCTTTCTCACTGTGTCAGCCATGTTCACTACAGATAAGCTG	6107	Db	217613	CTTCACAAATACCTTAATGAAACCTTAATCCAGAGGSGCTGTCTGATG-CCAG 217555
Db	218661	ATTTGCTCGTTTACT---CATGATGTCACAGCGGGCTACCAACATAGCTGA	218604	QY	7181	ATACAGATTAAATTTATGTTGCTTAATTAAATAAGTACTGAGCTTCTGTT 7180
QY	6108	TGTGATCACATAGATCACATAAGCTGCACTGTTCTGAACTAGCTG	6167	Db	217554	ACTCCAACACTACATCACTACTGCCACT-----TCIACCAAGCTTCTGG 217505
Db	218603	T-TGTGTAAGCTCCGTAATACCTTAATGTTATGT	218552	QY	7241	GCCTACAGTTGTTGTTGCTGAAATAAGATATGGAATACAAAMATAATGAAAT 7300
QY	6168	CTTGGGATCATGAAAGAGATCTGGACCACATCTGGAAAGCACACTCTGATG	6227	Db	217504	ACCTGCAGTGTGTTGTTGTTGTTGATGAAAGATTATGAAATGCCCACAG 217446
Db	218551	CTCTGAGATTAAGCACTGTTGCTTCTGAACTGCTGAACTGCTG	218500	QY	7301	AAAAAT 7305
QY	6228	CACTTGGCTTCTGATGTTGCACTGAACTGAACTGTTATGTTGCTTGTG	6287	Db	217445	AGAT 217441
Db	218499	CCAATA-AGCCTTCTCAATGTTGAGGCC-CTGGAGTGGAGGCCAGTGT	218322	RESULT 8		
QY	6288	T-GTTATCTAGC-CTAGAAATGCAAGAACACTTGTCTATGTTAAACACTCTTA	6344	AF05736_6		
Db	218441	TAAGTTCTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC	218382	LOCUS	AF057366	1575 bp rRNA linear
QY	6345	TATAAGCTTCAGTGTCACTGAACTGAACTGAACTGTTATCTCTCTGAAACACTCT	6404	DEFINITION	Mus musculus C2H2 zinc finger protein PLAG1 (Plag1)	mRNA, complete
Db	218381	TATAAGCTTCAGTGTCACTGAACTGAACTGAACTGTTATCTCTGAAACCTCT	218322	ACCESSION	AF057366	cds.
QY	6405	TCAATATTATGGCATGCTGTAATGTAACCTCTCATGGCTAACACTCTCATGGCTA	6464	VERSION	AF057366-1	GI:6650194
Db	218321	TCAACGTTCTATGGGATCATCTGTCATGTCATGGCTAACACTCTCATGGCTA	218262	KEYWORDS	Mus musculus (house mouse)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi;					

REFERENCE	AUTHORS	JOURNAL	FEATURES	BASE COUNT	ORIGIN
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1575)	Kas, K., Voz, M.L., Roijer, B., Aström, A.K., Moyen, E., Stenman, G. and Van de Ven, W.J.	Nat. Genet. 15 (2), 170-174 (1997)	Db Qy	361 881	ATGACCCCAAAAGAACCTTAAATGGAGAGGTGGCAAGAGTACAATCACAGC 4.20
Promoter swapping between the genes for a novel zinc finger protein and beta catenin in pleiomorphic adenomas with t(3;8) (p21;q2)	Hersens, K., Voz, M.L., Van de Ven, W.J.M. and Kas, K.	Direct Submission	Db Qy	421 941	TTCGATTAAACGTCACCTGGCCTTCATGCCGCAACAAAGCGTGAACCTCACCTGTAAGG 9.40
Submitted (06-APR-1998) Laboratory of Molecular Oncology, Center of Human Genetics, Herestraat 49, Leuven 3000, Belgium	Location/Qualifiers	1..1575	Db Qy	481 1061	TTCGTTGGAGAAATTGGAGCACCTGCTCTAGACCACTGAAATCTCACCGG 4.80
2 (bases 1 to 1575)	Hersens, K., Voz, M.L., Van de Ven, W.J.M. and Kas, K.	Submitted (06-APR-1998) Laboratory of Molecular Oncology, Center of Human Genetics, Herestraat 49, Leuven 3000, Belgium	Db Qy	1001 541	GCAAGTGTCGAACCTTGAAGCAGGGAGCTTGTGACACCTTAACCTCATGCAG 1.000
REFARENCE	AUTHORS	JOURNAL	source	1..1575	TATGTTGCGAACCTTGAAGCAGGGAGCTTGTGACACCTTAACCTCATGCAG 1.060
MEDLINE			/organism="Mus musculus"	/note="synonym: pleiomorphic adenoma gene 1"	TGTTGCGAACCTTGAAGCAGGGAGCTTGTGACACCTTAACCTCATGCAG 6.00
PUBMED			/mol_type="mRNA"		TCTACACCCGAAAGGATGTCGGAGACATGGTGGCACACTGGAGAAGGACTTCC 5.40
9030842			/strain="129/SV"		TCTACACCCGAAAGGATGTCGGAGACATGGTGGCACACTGGAGAAGGACTTCC 5.40
			/db_xref="taxon:10090"		TCTGTCAGTATTGCGAACAGATTGGCAAAGGATCACCTGATCGACATGAGA 1.180
			/map="t (3;8) (p21;q2)"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/tissue type="pleiomorphic adenoma"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			gene	1..1575	TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/gene="Plag1"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/note="synonym: pleiomorphic adenoma gene 1"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			21..1520		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/gene="Plag1"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/codon_start=1		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/product="C2H2 zinc finger protein PLAG1"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/protein_id="RAF2162.1"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/db_xref="GI:1650195"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			/translation="MATVPGDLSSEVRDTOKAAPSGRKRIGESKPKRKNPQCLDKAFN SVEKLKVHSFHSGERPYKTHQDCTKAFTSKYLUQRIMATHSPPEKTHKCNYCKRMFH RKDHLLKHJLHTHDNPKECEBCECECSYNTKLQGRHLAHLUHARQHGRKDELCQYC AQRGRKDILTRMKKSHNBLLKVTEPVDLDFTCNNMSVPKDELPVMSLSSCB LLSPPFTNTQLQNYNTPOMSGSAHOMITLPLGMTCPIUDNDAVPHSHHIAFKC PFSTSYATSYATSYATPEOPLKGEIESYMLQGGAQSSQDPSASSSKLGLBQSSGPDD GAGDLISLKSISIISDPLSTPALDSQLENFIPUNGPPYNTLVSQSLGMSYSQEEBHS SVSLOPLQTODLQDANTVGLSSHLSAFTSSLSSSTLPRHQAFQ"		TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			436 a	417 c	TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			347 g	375 t	TCTGTCAGTACTGCGAACAGAGTTGGCAAAGGATCACCTGACATGAGA 7.20
			Query Match	17..77;	Score 1295.6; DB 10; Length 1575;
			Best Local Similarity	89.6%;	Pred. No. 1.4e-220;
			Matches 1416;	0; Mismatches 159;	Indels 5; Gaps 2;
			Qy	461 1	AAGGCTCTGGTTAGGTGGCTGGCCACTGTGATTCCTGGTATTGTGAGTAGAG 520
			Dbb	521	ATACCCAAGGTCCTTCAGGGAAAGTAAGCTGTTGTAACCAACCAAAACT 580
			Qy	61	ATACCCAAGGTCCTTCAGGGAAAGTAAGCTGTTGTAACCAACCAAAACT 120
			Dbb	581	TTCCTTGCCAACTGTGTGACAGGCCCTTAACAGTGTGAAATTAAGGTTCACTCCT 640
			Qy	121	TTCCTTGCCAACTGTGTGACAGGCCCTTAACAGTGTGAAATTAAGGTTCACTCCT 180
			Dbb	641	ACTCTCACACGGAGAGGGCTCAAGTGTGACAGGCCCTTAACAGTGTGAAACCAAAACT 760
			Qy	181	TTCCTTGCCAACTGTGTGACAGGCCCTTAACAGTGTGAAACCAAAACT 300
			Dbb	701	TTCCTPAAGTCAAATTACAAGGGCATGGCTCAAGTGTGAAACCAAAACT 760
			Qy	241	TTCCTPAAGTCAAATTACAAGGGCATGGCTCAAGTGTGAAACCAAAACT 300
			Dbb	761	GTAATTATTGTGAGAAAATGTGTCACCGGAAGATCATCGAGAAATCACCTCCATACAC 820
			Qy	301	GTAATTACTGTGAGAAAATGTGTCACCGGAAGACCATCACCTCCATACAC 360
			Qy	821	ACGACCTAAAGAAGACGTTAAAGTGCGAGAAATGTGGAGAAACTACATAACAGC 880
				1901	CCTCTGCACTCAGTGTGAGGTTCAACAGGACTTGTGAGGAACTACCCCAAC 1960

Db	1438	CTCTGCACCTCACTGTCAGCAGCGCTTACCCAGGCCCTAGCTCAAGGACTAACCTGGCCC	1497	
Qy	1961	GTTTCCATCAAGCTTTCAGTAGATTTAGATGTTGGACATGGATTTAGAAGATGTATGTC	2020	
Db	1498	GTTTCCATCAAGCTTTCAGTAGATTTAGATGTTGGCCGAGGAACTCA--CCCGAGGTGCTTG	1555	
Qy	2021	TAGTGTCGCCCTAGATGACC	2040	
Db	1556	TAGCCCTGCTGATGATGRC	1575	
RESULT	9			
LOCUS	AC090187	65748 bp DNA linear	HTG 13-APR-2001	
DEFINITION	Homo sapiens chromosome 8 clone RP11-140116 map 8, LOW PASS			
SEQUENCE SAMPLING.				
VERSION	AC090187			
KEYWORDS	GI:13621271			
HTG;	HRGS PHASED.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1 (bases 1 to 65748)				
Birren,B., Linton,L., Nusbaum,C. and Lander,E.				
Homo sapiens chromosome 8, clone RP11-140116				
JOURNAL	Unpublished			
REPERIENCE	2 (bases 1 to 65748)			
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barua,N., Bastion,V., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Faro,S., Cooke,P., DeArellano,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goettte,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horcon,L., Hulme,W., Iliev,I., Johnson,R., Janes,C., Karafatos,A., LaRocque,K., Lamazares,R., Landers,T., Lehozcy,J., Levine,R., Liu,G., Maclean,C., McDonald,P., Marquis,N., Matthews,C., McCaffery,M., McEwan,P., McKernan,K., McPheeers,R., Meldrim,J., Mewcaus,L., Mihova,T., Murphy,T., Nayor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunthang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Riss,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severay,P., Sognez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zatouni,J., Zembek,L., Zimmer,A. and Zody,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141 USA			
COMMENT	All repeats were identified using RepeatMasker: Smit, A.P.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html			
Center	Whitehead Institute / MIT Center for Genome Research			
Center code:	WIBR			
Web site:	http://www-seq.wi.mit.edu			
Contact:	sequence_submissions@genome.wi.mit.edu			
----- Project Information				
Center project name:	L12184			
Center clone name:	140_I_16			
----- Genome Center				
Center:	Whitehead Institute / MIT Center for Genome Research			
On Apr 13, 2001 this sequence version replaced gi:13273404.				

* NOTE: This record contains 78 individual				
* sequencing reads that have not been assembled into				
* contigs. Runs of N are used to separate the reads				
* and the order in which they appear is completely				
* arbitrary. Low-pass sequence sampling is useful for				
* identifying clones that may be gene-rich and allows				
* overlap relationships among clones to be deduced.				
* However, it should not be assumed that this clone				

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barnes,N., Bastien,V., Boguski,M., Boukhalter,A., Brown,A., Camarata,J., Campopiano,A., Choeper,Y., Colangeli,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardya,S., Ginde,S., Goyetta,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Huime,W., Iliev,I.I., Johnson,R., Jones,C., Karatas,A., Latocque,K., Lamazaress,R., Landers,T., Lehoocky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meineus,L., Mihova,T., Milenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rose,C., Rogov,P., Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schuppach,R., Seaman,S., Sevary,P., Sougnez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talama,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Vieil,R., Vo,A., Wilson,B., Wu,X., Wuman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (17-FBB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141 USA

COMMENT On Apr 13, 2001 this sequence version replaced gi:13273404.

REMARKS All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute / MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@gnome.wi.mit.edu
Project Information
Center project name: Li21B4
Center clone name: 140_I_16

NOTE: This record contains 78 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

1 742: contig of 742 bp in length
* 743 842: gap of 100 bp
* 843 1608: contig of 766 bp in length
* 1609 1708: gap of 100 bp
* 1709 2435: contig of 727 bp in length
* 2436 2535: gap of 100 bp
* 2536 3238: contig of 703 bp in length
* 3239 3338: gap of 100 bp
* 3339 4054: contig of 716 bp in length
* 4055 4154: gap of 100 bp
* 4155 4896: contig of 742 bp in length
* 4897 4996: gap of 100 bp
* 4997 5757: contig of 761 bp in length
* 5758 5857: gap of 100 bp
* 5858 6611: contig of 759 bp in length
* 6617 6716: gap of 100 bp
* 6717 7489: contig of 773 bp in length
* 7490 7589: gap of 100 bp
* 7590 8354: contig of 765 bp in length
* 8355 8454: gap of 100 bp
* 8455 9217: contig of 763 bp in length
* 9218 9317: gap of 100 bp
* 9318 10039: contig of 722 bp in length
* 10040 10134: gap of 100 bp
* 10140 10863: contig of 729 bp in length

* 10869 10968: gap of 100 bp
* 10969 11700: contig of 732 bp in length
* 11700 11800: gap of 100 bp
* 11801 12553: contig of 753 bp in length
* 12554 12653: gap of 100 bp
* 12654 13439: contig of 786 bp in length
* 13440 13539: gap of 100 bp
* 13540 14304: contig of 765 bp in length
* 14305 14404: gap of 100 bp
* 14405 15152: contig of 748 bp in length
* 15153 15252: gap of 100 bp
* 15253 16016: contig of 764 bp in length
* 16017 16116: gap of 100 bp
* 16117 16662: contig of 746 bp in length
* 16663 16962: gap of 100 bp
* 16963 17685: contig of 723 bp in length
* 17686 17785: gap of 100 bp
* 17786 18512: contig of 727 bp in length
* 18513 18612: gap of 100 bp
* 18613 19341: contig of 729 bp in length
* 19342 19441: gap of 100 bp
* 19442 20175: contig of 734 bp in length
* 20176 20275: gap of 100 bp
* 20276 21014: contig of 739 bp in length
* 21015 21114: gap of 100 bp
* 21115 21879: contig of 765 bp in length
* 21880 21979: gap of 100 bp
* 21980 22736: contig of 757 bp in length
* 22737 22836: gap of 100 bp
* 22837 23616: contig of 780 bp in length
* 23617 23716: gap of 100 bp
* 23717 24415: contig of 699 bp in length
* 24416 24516: gap of 100 bp
* 24516 25285: contig of 770 bp in length
* 25286 25385: gap of 100 bp
* 25386 26159: contig of 774 bp in length
* 26160 26259: gap of 100 bp
* 26260 27032: contig of 773 bp in length
* 27033 27852: contig of 720 bp in length
* 27853 27952: gap of 100 bp
* 27953 28686: contig of 734 bp in length
* 28687 28786: gap of 100 bp
* 28787 29508: contig of 722 bp in length
* 29509 29608: gap of 100 bp
* 29609 30359: contig of 751 bp in length
* 30360 30459: gap of 100 bp
* 30460 31206: contig of 747 bp in length
* 31207 31306: gap of 100 bp
* 31307 32050: contig of 744 bp in length
* 32051 32151: gap of 100 bp
* 32151 32913: contig of 762 bp in length
* 32913 33012: gap of 100 bp
* 33013 33739: contig of 727 bp in length
* 33739 33839: gap of 100 bp
* 33840 34380: contig of 741 bp in length
* 34380 34480: gap of 100 bp
* 34480 35313: contig of 693 bp in length
* 35313 35374: gap of 100 bp
* 35374 35473: contig of 718 bp in length
* 35473 36191: contig of 757 bp in length
* 36192 36292: gap of 100 bp
* 36292 37022: contig of 731 bp in length
* 37023 37122: gap of 100 bp
* 37122 3774: contig of 752 bp in length
* 3774 37975: gap of 100 bp
* 37975 38731: contig of 757 bp in length
* 38731 38831: gap of 100 bp
* 38831 39605: contig of 774 bp in length
* 39605 39705: gap of 100 bp
* 39705 40465: contig of 760 bp in length
* 40466 40565: gap of 100 bp
* 40566 41271: contig of 706 bp in length
* 41271 41371: gap of 100 bp

Query Match 42138: contig of 767 bp in length										
Best Local Similarity 8.9%; Score 651.8; DB 2; Length 65748;										
Matches 695; Conservative 0; Mismatches 122; Indels 3; Gaps 3;										
*	41372	42138:	contig of 767 bp in length							
*	42139	42238:	contig of 100 bp							
*	42239	42959:	contig of 721 bp in length							
*	42960	43059:	gap of 100 bp							
*	43060	43786:	contig of 727 bp in length							
*	43787	43886:	gap of 100 bp							
*	43887	44623:	contig of 737 bp in length							
*	44624	44723:	gap of 100 bp							
*	44724	45463:	contig of 740 bp in length							
*	45464	45563:	gap of 100 bp							
*	46273:	46373:	contig of 710 bp in length							
*	46374	47119:	gap of 100 bp							
*	47120	47219:	contig of 746 bp in length							
*	47220	47974:	gap of 100 bp							
*	47975	48074:	gap of 100 bp							
*	48075	48802:	contig of 728 bp in length							
*	48803	48902:	gap of 100 bp							
*	48903	49672:	contig of 746 bp in length							
*	49673	49772:	gap of 100 bp							
*	49773	50545:	contig of 755 bp in length							
*	50546	50645:	gap of 100 bp							
*	51367:	51467:	contig of 722 bp in length							
*	51368	51468:	gap of 100 bp							
*	51468	52183:	contig of 716 bp in length							
*	52184	52283:	gap of 100 bp							
*	52284	53020:	contig of 737 bp in length							
*	53021	53120:	gap of 100 bp							
*	53121	53852:	contig of 732 bp in length							
*	53853	53952:	gap of 100 bp							
*	53953	54700:	contig of 748 bp in length							
*	54701	54800:	gap of 100 bp							
*	54801	55550:	contig of 750 bp in length							
*	55551	55650:	gap of 100 bp							
*	55651	56510:	contig of 768 bp in length							
*	56510	56518:	gap of 100 bp							
*	56519	57273:	contig of 755 bp in length							
*	57273	58093:	gap of 100 bp							
*	58093	57374:	contig of 720 bp in length							
RESULT 11										
	AC107961	AC107961	49357 bp	DNA	linear	RP11-1053A18	map 8,	HTG	24-JAN-2002	LOW-PASS
			Homo sapiens chromosome 8 clone							
			SEQUENCE SAMPLING.							
			AC107961.1	GI:18108689						
			HTG; HTSS PHASE0.							
			Homo sapiens (human)							
			Homo sapiens							
			Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;							
			Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo							
			RESULTS 1							
	AC107961	AC107961	49357 bp	DNA	linear	RP11-1053A18	map 8,	HTG	24-JAN-2002	LOW-PASS
			Homo sapiens chromosome 8 clone							
			SEQUENCE SAMPLING.							
			AC107961.1	GI:18108689						
			HTG; HTSS PHASE0.							
			Homo sapiens (human)							
			Homo sapiens							
			Birren,B., Linton,L., Nelson,C., Allen,N., Lander,E., Birren,B., Linton,L., Nelson,C., Allen,N., Lander,E., Anderson,S., Barna,N., Bastien,V., Boguslavsky,I., Boukhgalter,B., Brown,A., Camarata,J., Campobiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangeiro,M., Collins,S., Collymore,A., Cooke,A., DeArellano,K., Dewart,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyetty,M., Graham,L., Grand-Pierre,N., Hagos,V., Horton,L., Iliev,I., Johnson,R., Jones,C., Kamat,A., Kortats,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., McDonald,D.P., Major,J., Margolis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Medrano,J., Menuez,L., Milhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spangler,B., Stange-Thomann,A., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.							
			REFERENCE							
			1 (bases 1 to 49357)							
			AUTHORS							
			Birren,B., Linton,L., Nelson,C., Allen,N., Lander,E.							
			TITLE							
			Unpublished							
			JOURNAL							
			2 (bases 1 to 49357)							
			REFERENCES							
			Birren,B., Linton,L., Nelson,C., Allen,N., Lander,E., Anderson,S., Barna,N., Bastien,V., Boguslavsky,I., Boukhgalter,B., Brown,A., Camarata,J., Campobiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangeiro,M., Collins,S., Collymore,A., Cooke,A., DeArellano,K., Dewart,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyetty,M., Graham,L., Grand-Pierre,N., Hagos,V., Horton,L., Iliev,I., Johnson,R., Jones,C., Kamat,A., Kortats,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., McDonald,D.P., Major,J., Margolis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Medrano,J., Menuez,L., Milhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spangler,B., Stange-Thomann,A., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.							
			COMMENT							
			Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA							
			All repeats were identified using RepeatMasker:							
			Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html							

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24594
 Center clone name: 1053_A_18

* NOTE: This record contains 62 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

1 692: contig of 692 bp in length
 * 693 792: gap of 100 bp
 * 1456: contig of 664 bp in length
 * 1556: gap of 100 bp
 * 2165: contig of 609 bp in length
 * 2166: gap of 100 bp
 * 2265: contig of 698 bp in length
 * 2963: contig of 698 bp in length
 * 2964 3063: gap of 100 bp
 * 3064 3778: contig of 715 bp in length
 * 3779 3878: gap of 100 bp
 * 3879 4538: contig of 660 bp in length
 * 4539 4638: gap of 100 bp
 * 4639 5277: contig of 639 bp in length
 * 5278 5378: gap of 100 bp
 * 6099: contig of 722 bp in length
 * 6100 6199: gap of 100 bp
 * 6200 6892: contig of 693 bp in length
 * 6893 6992: gap of 100 bp
 * 6993 7707: contig of 715 bp in length
 * 7708 7808: gap of 100 bp
 * 8515: contig of 708 bp in length
 * 8615 8616: gap of 100 bp
 * 8616 9299: contig of 684 bp in length
 * 9300 9399: gap of 100 bp
 * 9400 10101: contig of 702 bp in length
 * 10102 10202: gap of 100 bp
 * 10896: contig of 695 bp in length
 * 10997 10996: gap of 100 bp
 * 11661 11760: contig of 664 bp in length
 * 11761 12438: gap of 100 bp
 * 12439 12558: contig of 702 bp in length
 * 12539 13233: gap of 100 bp
 * 13234 13333: contig of 724 bp in length
 * 13334 14057: gap of 100 bp
 * 14058 14158: contig of 678 bp in length
 * 14158 14838: contig of 681 bp in length
 * 14839 14938: gap of 100 bp
 * 14939 15642: contig of 695 bp in length
 * 15643 15742: gap of 100 bp
 * 15743 16661: contig of 719 bp in length
 * 16462 16561: gap of 100 bp
 * 16562 17250: contig of 689 bp in length
 * 17251 17351: gap of 100 bp
 * 18053 18153: contig of 704 bp in length
 * 18054 18862: gap of 100 bp
 * 18154 18863: contig of 709 bp in length
 * 18863 19683: gap of 100 bp
 * 19684 20488: contig of 705 bp in length
 * 20489 20589: gap of 100 bp
 * 20589 21312: contig of 724 bp in length

FEATURES
 source

Location/Qualifiers
 1. .49357
 /organism="Homo sapiens"

		BASE COUNT ORIGIN	399 a	577 c	493 g	360 t
Qy	4935 ATTCTCTAGTGGCCAGATCATTATCCCTGAATTTCGGTAGTGCTTAGCTTGGTT 4992	Query Match Best Local Similarity Matches 826; Conservative 0; Mismatches 542; Indels 48; Gaps 6;				
Db	58 ATTCTCTAGTGGCCAGATCATTATCCCTGAATTTCGGTAGTGCTTAGCTTGGTT 1					
RESULT 13						
LOCUS	BC023655	1829 bp mRNA, linear	PRI 23-SEP-2002			
DEFINITION	Homo sapiens, clone MGC:29597 IMAGE:4764127, mRNA, complete cds.					
ACCESSION	BC023655					
VERSION	BC023655.1					
KEYWORDS	MGC.					
SOURCE	Homo sapiens (human)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1829)					
REFERENCE	Strausberg, R.					
AUTHORS						
TITLE	Submitted (05-FEB-2002) National Institutes of Health, Mammalian Direct Submission (MGC). Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
JOURNAL						
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov					
COMMENT	Contact: MGC help desk Email: cgbabs@nlm.nih.gov					
	Tissue Procurement: Louis Staudt					
	CDNA Library Preparation: Rubin Laboratory					
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)					
	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;					
	Web site: http://www.nisc.nih.gov/					
	Contact: nisc_mgc@nigri.nih.gov					
	Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Deitchman, N.L., Granite, S., Guan, J., Gupta, P., Haghghi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.R., Masiello, C., Mastriani, B., McCloskey, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsourelis, C., Voigt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.					
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at: http://image.liln.gov					
	Series: IRL Plate: 39 Row: f Column: 4					
	This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6031195.					
FEATURES	Location/Qualifiers					
source	1 . 1829					
	organism="Homo sapiens"					
	/mol_type="tRNA"					
	/db_xref="taxon:2606"					
	/clone=MGC:29597 IMAGE:4764127"					
	/tissue_type="Primary B-Cells from Tonsils"					
	/clone_lib="NIH_MGC_48"					
	/lab_host="DH10B_R"					
	/note="vector: pOTB7"					
CDS	2.13 . 1703					
	/codon_start=1					
	/product="Unknown (protein for MGC:29597)"					
	/protein_id="AAH23655.1"					
	/db_xref="GI:23271120"					
	/translation="MTCIFETSYVPPWILDAKQEEEVGKLVPRGRGEAESOYKCOCEI SGTPPSNGEKLRPHSLIQPQRDPSYSPQLHCKAFASKVKKLYHMATHSAQKPHQCMY CDRMFEKRHDHLNLHQHDPNKEALHSEKPCNDKLRFLYRHLKMDHAASQGLSKRD CLQTFESTOQALLEBLHKHSRIVAGGAKKEPKCDHCDRFLYRHLVHTGSKRD PLOQCAORGEKHDLTFLHVKSXSHSQUELLKTFEPUDMLGLISCSSTVSVREBELSPVLCMA SRDYMGTGTRAPPGMPLPMGQTAHPTMPGSPVPLVHNTLPGSLSLSSAROPASPOPAAAULID EALLAKSPANLSEALCANVUDSHLIGFLPLNPUPCNPGATGGLUNGYSQAEAQPLI TTQAQPOQSPGAGLNPFGPLSLPVFTSGLSLSTLPRHQAFQ"					

Db	1272 TTGCCAAAGTGGAGGTGGATAGTTCTGGGGAGCTTCTGGAGCCNGTCCTCTCA 1331	misc_feature	/clone lib="HEP"
QY	1591 TCCCCAGATCTCAGCATCTCATTAAGGCTATCTCCAGATTGGTCC 1650		/note="cloning vector pME18SPL3"
Db	1332 TCCGCTGAACCCAGCCGGCTCA-----CTTAGCCGGCA 1370		1. .2429
QY	1651 CTAGATGATGGCGGGAGACCTCTCCATTCCATCAGTGACCCC 1710		/note="highly similar to Aff06005 Homo sapiens zinc finger protein PLAGL2 mRNA"
Db	1371 GCTGGGCCCTCCATGATGAGACTGCTGGCCAGGGCCCAACCTCTCTGAGGCC 1410	BASE COUNT	540 a
QY	1711 CTAACACACCAGCATGGATTTCCTCAGTTAAATTCTCATACCTTTAATGGTCT 1770	ORIGIN	743 c
Db	1431 CTCRGCGCTTAATGTGGACTCTCCCACCTACTGGGTTCTCCACCTCARCTGCC 1490	Query Match	5.4%
QY	1771 CCCTATAA--TCCCTCTATCAGTGGGGCTGGCTACTCCAGGTAAAGCTATCCAGGAAAGCA 1827	Best Local Similarity	58.3%
Db	1491 CGGTAAACCCACCTGGGGCACAGGGGGCTGGTCATGGCTGAGGA 1550	Matches	826;
QY	1828 CATCTCTGTTCCAGGCTCCACACACAGGATCTCAGGATCTGCAAAACACT 1887	Conservative	0;
Db	1551 CAGGCCCTGCTTACCTTGTGCAAGCTCGCCCTAGATTCAGGAGGACA 1610	Mismatches	542;
QY	1888 ATAGGGCTTGGCTCTGGACTCACTGTAGCAGCTTACCCAGGAGTTAACGACAGCT 1947	Indels	48;
Db	1611 CTGACTTGGCTCTGGACTCTGGCTCTCTAACCTGTGGCTGACT---AGC 1667	Gaps	6;
QY	1948 ACCACCCPCCCAAGTTCCATCAAGCTTTCAGTAG 1983		
Db	1668 ACCACCCCTCCCTCTCCATCAAGCATCCAGTAG 1703		
RESULT 14			
AK026936	AK026936 Homo sapiens cDNA: FLJ23283 2429 bp mRNA linear PRI 29-SEP-2000	DEFINITION	
	AK026005 Homo sapiens zinc finger protein PLAGL2 mRNA.	ACCESSION	
VERSION	AK026936.1 GI:10439912	VERSION	
KEYWORDS	Oligo capping; fis (full insert sequence).	KEYWORDS	
ORGANISM	Homo sapiens (human)	ORGANISM	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 (sites)	Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikeda,Y., Okano,S., Okutani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibaahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	REFERENCE	
AUTHORS	NEDO human cDNA sequencing project	AUTHORS	
JOURNAL	2 (bases 1 to 2429)	JOURNAL	
REFERENCE	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibaahara,T., Tanaka,T. and Nakamura,Y.	REFERENCE	
AUTHORS	Direct Submission	AUTHORS	
TITLE	Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108 8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,	JOURNAL	
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).	COMMENT	
FEATURES source	Location/Qualifiers	FEATURES source	
	1. .2429		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="HEP0729"		
	/cell_line="HePG2"		
	/cell_type="hepatoma"		

RESULT 15
 AF006005 AF006005 4051 bp mRNA linear PRI 07-MAR-2000
 DEFINITION Homo sapiens zinc finger protein PLAGL2 mRNA, complete cds.
 ACCESSION AF006005
 VERSION AF006005.1 GI:3513455
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
 REFERENCE 1. (bases 1 to 4051)
 AUTHORS Kas, K., Voz, M.L., Hensen, K., Meyen, E. and Van de Ven, W.J.M.
 TITLE Transcriptional activation capacity of the novel PLAG family of
 zinc finger proteins
 JOURNAL J. Biol. Chem. 273 (36), 23026-23032 (1998)
 MEDLINE 972527
 PUBMED 98389728
 REFERENCE 2. (bases 1 to 4051)
 AUTHORS Kas, K., Hensen, K., Meyen, E., Voz, M.L. and Van de Ven, W.J.M.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-1997) Laboratory for Molecular Oncology, Center
 of Human Genetics, Herestraat 49, Leuven 3000, Belgium
 Location/Qualifiers 1..4051
 FEATURES source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="mRNA"
 /taxon: 9606
 /chromosome="20"
 /map="20q11"
 1..4051
 gene PLAGL2
 /gene="PLAGL2"
 /geno="PLAGL2"
 /note="C2B2-type zinc finger protein; PLAG-like 2"
 /codon_start=1

/product="zinc finger protein PLAGL2"
 /protein_id="AAC34252.1"
 /db_xref="GI:3513456"
 /translation="MVTTEPFSVPWFLQDQKEEVEVGMKLVPRPGREAEASQVKCCOEI
 SGTPPSNGEKLRPHSLQPQRPIYPSQPIHEKGAKAFSKYKLYPREMATHAQKPHQCMY
 CDKMEFHKDHLRNHLQTHDNKEALHCSSEGKNTNLGYRRHLAMHAASSGDLSCKV
 CLQFTESTSQQALEHKAHSRVRVAGAKEKKHPDCDRETRYEDVRLHVVHTSRKD
 FLCQYKAQRGRKDKHTRAYXKQKLSQSLTMLKTEPVBLIGLSSAQSSTSYKEVSPVL
 CMASPRVYMGTKAFPGLMOMYGAHTPTMVEVPSLSSAEPQASPOPAAAAAILD
 PAQLPKPYQIGSTSVLIDPKLPVEDSFELAELPGLSLSAEPLNPKNPVATGTLPMGSYBLESPLS
 EALLAKSPANIEALCAANYDIFSHLFLGPFLNPKNPVFTGSLSSLTLPREHFQAFQ"
 BASE COUNT 874 a 1192 c 972 g 1013 t
 ORIGIN
 Query Match 5.4% Score 392.8; DB 9; Length 4051;
 Best Local Similarity 58.3%; Pred. No. 6..8e-60;
 Matches 826; Conservative 0; Mismatches 542; Indels 48; Gaps 6;
 Matches 826
 Qy 577 AACTTTCTTGCAACCTGTGACAAGGCCCTTAACAGTGTGAGAAATTAAAGGTTCAAC 636
 Db 317 AAGGCCAATGTGAAATTTCGGGAACACCTTCTCAATGGGAAAGAAGCTGACGCC 376
 Qy 637 TCCTAGCTCACACAGGAGAGCCCTAACAGTGTGACAAAGTGCACAAAGGCC 696
 Db 377 AGCTCCCGAACACAGAGACGAGACCATATAGTGCCTCTAGTGCACAGGCT 436
 Qy 697 TTGTTCTTAAGTACACAATTACAAAGGACACATGCTACTCATCTCCAGAAACCCAC 756
 Db 437 TTCTCTCCAATAAACAGTGTATAGCACATGCCAACACTAGGCCAGAACCCAC 496
 Qy 757 AAGTGAATTATTGTTGAGAAACCTTCACCGGAAAGATCATCTGAAATACTCTCAT 816
 Db 497 CAGGTATGTTACTGTGATAGATTTAACAGGAAAGTGTGCAAGAACTAAATACC 556
 Qy 817 ACACAGACCCCTAACAAAGAGCCTTAACAGTGTGAAAGATGTGCAAGAACTAAATACC 876
 Db 557 ACCATGATCTCAACAAAGGCCCTCCACTGCTGAGTAAAGCTTAAAGTAAATAGATTAAATACG 616
 Qy 877 AAGGTTGGGTTAACGTGACTTGCCTGTCATGCCGAAACAATGGTGAACCTTCTGT 936
 Db 617 AAGCTGGGTPACCGGCCACCTGCCATGATGTGCCAGGGGACCTCTGGCTGC 676
 Qy 937 AAGGTATGTTGCAACATTGTTGAAGGACCCCTAAATCTCAT 996
 Db 677 AAGGTGTGCTGCAAGACCTTGTGAAAGTACCTGCTGTCATGCCGAAACCCAC 736
 Qy 997 GCAGGGAAAGTCGTCTGGGGTTAACAGAAAGAACAGTGGGAGCTCTGGTATCGC 1056
 Db 737 TCACGCCGGTAGGGCTGCAAGGGAAAGAACAGCCTGTGACCTGTGGTAAAGGAC 796
 Qy 1057 CGGTRCTACCCGAAGGATGTGGAGAACAGTGGTGTGAAAGAACAGGAC 1116
 Db 797 CGGTRCTATACTGTAAAGGTGTAAGGTGACGGGCACACTAGTGTGGTAAAGGAC 856
 Qy 1117 TTCTCTGTCAGTATGTCAGAGATTGGGGAAAGGATCACTGTACTGACATATG 1176
 Db 857 TTCTCTGCGGATCTGTGCTGTCAGTGTGGCTAACAGAACAGTGGGATTCCTGAC 916
 Qy 1177 AAGAAGAGTCACATAGGGCTCTGAAGTAAACACAAACAGTGGGATTCCTGAC 1236
 Db 917 AAGAAGAGCCACTGGAGCTGGCTCAAGACAGGCCCTGGACATGTAGGGC 976
 Qy 1237 CCATTACCTGCAATGTGCTGTCAGTGTGGCTAACAGAACAGTGGGATTCCTGAC 1296
 Db 977 CTACTCTGTCAGTGTGGCTAACAGAACAGTGGGATTCCTGAC 1036
 Qy 1297 CCTTCAGTGA -- ACTGTGTTACAACTAGGGCTCTGAAGTAAACACTCTAC 1353
 1..4051
 gene PLAGL2
 /gene="PLAGL2"
 CDS 203..1693
 /note="C2B2-type zinc finger protein; PLAG-like 2"
 /codon_start=1

Db	1097	GGTCCCCACATCCCTTACCATGCCAGCACGGCGTGCACACTCCCTGGTGACACACGG	1156
Qy	1411	TTACCCATTGGAATGACATGCCAATAGATAATGGAACTGACTGTTCACTCCCTTCACACCTT	1470
Db	1157	CTGCCCATGGGATGAGCTAACCTCTCGTAATCTACCTTCTCGTAATCTCAGCTAGCTC	1216
Qy	1471	TCTTCAAAATCCGPTCAAGTCTAACCTCATGCAATTCTATCTGAAAAGAACAG	1530
Db	1217	CCTCCAAAATACCAAGCTTGAATCTACCTCATGCAATTCTATCTGAAAAGAACAG	1261
Qy	1531	CAATTAAAGGGAAATTGAGGTAACTGATGGACTAACAGTGGCTTCAGTGGCTTCAGCTTC	1590
Db	1262	TGCCCCAAGCTGGAGGTGGATAGTTTCTGGGGACTTCCTGGAAAGCTGCTCTCTCA	1321
Qy	1591	TCCCAAGATCTCAAGGATCTCATCATCTAAGCTAGGTGGATCCTCAAGATGGTCC	1650
Db	1322	TCCGCTAACCCAGCCGGCTCA-----	1360
Qy	1651	CTAGATGATGGTGCAGGAGACCTCCCTATCCATCAAAGCTCTATCTCATAGTAACTCC	1710
Db	1361	GCTGCGCCCTCTCTAGTGAAGCACTGCTGGCAAGGCGCCAAACCTCTGAAGGCC	1420
Qy	1711	CTAAACACCCAGCATGGATTTCAGTGTATAATTCATACTTAAATGGTCCCT	1770
Db	1421	CTCTGGGTGTTAATGGACTCTCCACCTACTGGCTTCTCCACTAACCTGCCC	1480
Qy	1771	CCCTATAA---TCCCTCTATAGTGGGAGCCTTGGATGAGCTATCCCGGAAGAAGCA	1827
Db	1481	CGGTGTAACCACTGGCCACAGGGCTGGCATGGCTACTCCAGCTGGCAAGGCC	1540
Qy	1828	CATTCTCTGTTCCAGCTCCACACAAACACGGATCTTCAGGATCTGCAAAACACT	1887
Db	1541	CAGCCCCCTGCTTACCACTTCAAGTCAGCTCAAGATCCCCAGGAGTGGGGACCA	1600
Qy	1888	ATAGGGCTTGGGACTCACTGTCAGCTTCAAGCTTCAAGATCCCCAGGAGTGGGGACCA	1600
Db	1601	CTGAATTGGGCCCTCTGCACTCTGGCTTCTGCTCTGCTCTGCTGCTGCTGAGT--AGC	1657
Qy	1948	ACCAACCTCCACGTTCCATCAAGCTTCAAGCTTCAAGT	1983
Db	1658	ACCACCTGCTCGTTCCATCAAGCTTCAAGCTTCAAGT	1693

Search completed: January 31, 2004, 13:36:46
 Job time : 2557 secs



GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 M nucleic - nucleic search, using sw model
 am on: January 31, 2004, 03:59:43 i Search time 1682 Seconds
 (without alignments)
 11736.617 Million cell updates/sec
 title: US-09-242-772-116
 refref score: 7313
 sequence: 1 ggccggcatacactacaat.....tatgaataaaaaatctcggtgcc 7313
 scoring table: IDENTITY_NUC
 Gapop 10_0 , Gapext 1.0
 searched: 2552756 seqs, 1349719017 residues
 total number of hits satisfying chosen parameters: 5105512
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries
 database : N_GenSeq_19Jun03:
 1: /SIDS1/gcadata/genseq/genseq-emb1/NA1980/DAT: *
 2: /SIDS1/gcadata/genseq/genseq-emb1/NA1981/DAT: *
 3: /SIDS1/gcadata/genseq/genseq-emb1/NA1982/DAT: *
 4: /SIDS1/gcadata/genseq/genseq-emb1/NA1983/DAT: *
 5: /SIDS1/gcadata/genseq/genseq-emb1/NA1984/DAT: *
 6: /SIDS1/gcadata/genseq/genseq-emb1/NA1985/DAT: *
 7: /SIDS1/gcadata/genseq/genseq-emb1/NA1986/DAT: *
 8: /SIDS1/gcadata/genseq/genseq-emb1/NA1987/DAT: *
 Nucleotide sequenc
 Human breast cancer
 Human breast tumour
 Human foetal liver
 Probe #898 for gen
 Human brain express
 Human bone marrow
 Probe #926 for gen
 Probe #939 used to
 Probe #905 used to
 Human liver single
 Human genome-deriv
 Human foetal liver
 Probe #10823 for g
 Human brain express
 Human bone marrow
 Probe #10160 for g
 Probe #14113 used
 Probe #5922 used t
 Human liver single
 Human genome-deriv
 Haematopoietic cel
 Human immune syste
 DNA encoding zinc
 Human zinc finger
 Haematopoietic cel
 Human cDNA sequenc
 Tumour suppressor
 Haematopoietic cel
 Human immune syste
 DNA transcription
 Human cDNA clone (Human cDNA sequenc
 Human immune syste
 Human chemically trea

ATTACHMENT

Key CDS Location/Qualifi
48111983

PR	22-AUG-1996;	96EP-0202339.						
XX	KU LEUVEN RES & DEV.							
PA	(KULE-) UNIV GÖTEBORGS AB.							
PA	(UYGO-) HOLDINGBOLAGET AB.							
PA	(LEUT-) LEUVEN RES & DEV.							
XX	Kas KP,	Stenman KGD,	Van De Ven WJM,	Voz ML;				
XX	WPI; 1998-132252/13.	DR	P-PSDB;	AAW37948.				
XX	New tumourigenesis T-genes and proteins - useful for, e.g. preparing							
PT	PTT antibodies for clinically diagnosing cells having non-physiological							
PT	proliferative capacity such as lipoblastomas							
XX	PS	Fig 4; 71pp; English.						
XX	This is the nucleotide sequence of the human PLAG1 gene. It is a							
CC	tumourigenesis gene (T-gene), which is isolated in the form of PLAG1,							
CC	PLAG2, and CTNNB1 genes. Their proteins can be used as a starting							
CC	point for preparing antibodies for clinically/medically diagnosing							
CC	cells having a non-physiological proliferative capacity as compared							
CC	to wild type cells, where the former cells are selected from both							
CC	benign and malignant tumours, as well as leukaemia and lymphomas.							
CC	Derivatives of the T-gene are also used in the diagnosis and							
CC	preparation of therapeutic compositions for the treatment of cancers,							
CC	such as nucleic acid derivatives, and antibodies. The T-gene may be							
CC	used as a starting point for designing suitable expression-modulating							
CC	compounds or techniques for the treatment of non-physiological							
CC	proliferation phenomena in humans or animals. Expression inhibitors							
CC	of the T-gene can be used in the treatment of diseases involving							
CC	benign or malignant tumours.							
XX	Sequence 7313 BP; 2287 A; 1327 C; 1356 G; 2342 T; 1 other;							
	Query Match	100.0%;	Score 7312.6;	DB 19;	Length 7313;			
	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 1;	Indels 0;	Gaps 0;		
	Matches 7312;	Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	GGCAAGCATACTACATACATGGCTGCTGAAAGGGCTTAAGGAAACATTCCAGGCC	60					
Db	1	GGCAAGCATACTACATACATGGCTGCTGAAAGGGCTTAAGGAAACATTCCAGGCC	60					
QY	61	GCGCGTCGACGCCGAAATATGAGAAAATTTAGAATTGAAATTCCGGGGGGTGTAAG	120					
Db	61	GCGGGTCGACGCCGAAATATGAGAAAATTTAGAATTGAAATTCCGGGGGGTGTAAG	120					
QY	121	GCGCGGGACGGCGGGAGGGAGGTGTTAAAGGCCCGGCTTGCTCTGGTGCCTT	180					
Db	121	GCGCGGGACGGCGGGAGGGAGGTGTTAAAGGCCCGGCTTGCTCTGGTGCCTT	180					
QY	241	CTAGACTTGGACGGGCCCTCAGATTGGCAAAATGGAGGATTGGATTCACCTCT	300					
Db	241	CTAGACTTGGACGGGCCCTCAGATTGGCAAAATGGAGGATTGGATTCACCTCT	300					
QY	301	TCCACGAAGTCATGGGACTGGCTTAAGTCATGGCTTTCATCATGTAAT	360					
Db	301	TCCACGAAGTCATGGGACTGGCTTAAGTCATGGCTTTCATCATGTAAT	360					
QY	361	CAGTCCTTTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT	420					
Db	361	CAGTCCTTTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT	420					
QY	421	TCATTTAGTGGAGTGGCACTTACGACCATGAAACTTGGAGGATTGGATTCACCT	480					
Db	421	TCATTTAGTGGAGTGGCACTTACGACCATGAAACTTGGAGGATTGGATTCACCT	480					
QY	481	ATGGCCACTGTCACTTCGTTCTGCTCAGTCAGTCAGTCAGTCAGTCAGTC	540					
Db	481	ATGGCCACTGTCACTTCGTTCTGCTCAGTCAGTCAGTCAGTCAGTCAGTC	540					

Qy	1621	AAGCTAGGGTGGATCCTCGATTGGTCCTAGATGATGTTGCAAGACCTCTCCCTA	1680
D b	1621	AAGTAGGGTGGATCCTCGATTGGTCCTAGATGTTGCAAGACCTCTCCCTA	1680
Qy	1681	TCCAAAAGCTCATCTCACTCCATAGTCACCCCCCTAACACACAGATGGTTTCAG	1740
D b	1681	TCCAAAAGCTCATCTCACTCCATAGTCACCCCCCTAACACACAGATGGTTTCAG	1740
Qy	1741	TGTGTAATTCAACCTTAAATGTCCTCCPATAATCTCPATCACTGGGGAGCCPT	1800
D b	1741	TGTGTAATTCAACCTTAAATGTCCTCCPATAATCTCPATCACTGGGGAGCCPT	1800
Qy	1801	GGATGAGTATTCCAGGAAGGCAACATCTCTGTTCCCAGCTCCACACAACA	1860
D b	1801	GGATGAGTATTCCAGGAAGGCAACATCTCTGTTCCCAGCTCCACACAACA	1860
Qy	1861	CAGGATCCTCAGAACATCTCTGCAAAACACTATAGGGTTGGTCTGCACTCTCAGA	1920
D b	1861	CAGGATCCTCAGAACATCTCTGCAAAACACTATAGGGTTGGTCTGCACTCTCAGA	1920
Qy	1921	GCTTTCACCGAGCTTTAAGCACAGTAACACATATGGGTTGCTCTGCACTGTCAG	1980
D b	1921	GCTTTCACCGAGCTTTAAGCACAGTAACACATATGGGTTGCTCTGCACTGTCAG	1980
Qy	1981	TAGGATTCTGGGACATGGATTCATTACAGAAATGTTGCTGTTGCTCATGAC	2040
D b	1981	TAGGATTCTGGGACATGGATTCATTACAGAAATGTTGCTGTTGCTCATGAC	2040
Qy	2041	ATTTTATTATTTAGTGCCTACTTTAACAGATAAAATTCTGCTTTGTATAATAGA	2100
D b	2041	ATTTTATTATTTAGTGCCTACTTTAACAGATAAAATTCTGCTTTGTATAATAGA	2100
Qy	2101	ATTTCTATTAGCCGATTAATAGAAACTACGCTTTAACAGTTAACAGATGTC	2160
D b	2101	ATTTCTATTAGCCGATTAATAGAAACTACGCTTTAACAGTTAACAGATGTC	2160
Qy	2161	GTGTCAGTTAGTTACCTGGTATTGCTGATCCTGCAATTGTCACATTAA	2220
D b	2161	GTGTCAGTTAGTTACCTGGTATTGCTGATCCTGCAATTGTCACATTAA	2220
Qy	2221	AGACATTCTTCTTCATCACTGGAAAGCCATTATAGTGTAAATGCCA	2280
D b	2221	AGACATTCTTCTTCATCACTGGAAAGCCATTATAGTGTAAATGCCA	2280
Qy	2281	TTTCAAAATTACTTTAGCTTAAATTCTTCACTTGTAACTTGTAAATCTCACC	2340
D b	2281	TTTCAAAATTACTTTAGCTTAAATTCTTCACTTGTAACTTGTAAATCTCACC	2340
Qy	2341	TTTGACATCTGGCTCATTAACAAATTAGCTTACTAGCTTAAATTGTTG	2400
D b	2341	TTTGACATCTGGCTCATTAACAAATTAGCTTACTAGCTTAAATTGTTG	2400
Qy	2401	TGAATAACCTCAAGGTTTAATTTCTTACTAGCTTCAATTGATTAACTAGTGC	2460
D b	2401	TGAATAACCTCAAGGTTTAATTTCTTACTAGCTTCAATTGATTAACTAGTGC	2460
Qy	2461	TCAATGATTAAGTCCATTAGCTTCACTGGAAAGATAAAATTGTTGTTGATACACCTAA	2520
D b	2461	TCAATGATTAAGTCCATTAGCTTCACTGGAAAGATAAAATTGTTGTTGATACACCTAA	2520
Qy	2521	TTCAGATCATTTAGGAAACAAAACCTTAAACACAAAGCATACGTTAAATTGTTG	2580
D b	2521	TTCAGATCATTTAGGAAACAAAACCTTAAACACAAAGCATACGTTAAATTGTTG	2580
Qy	2581	TATGAGCCTTAAGGAAACAAAACCTTAAACACAAAGCATACGTTAAATTGTTG	2640
D b	2581	TATGAGCCTTAAGGAAACAAAACCTTAAACACAAAGCATACGTTAAATTGTTG	2640
Qy	2641	AGACTCTGTTAGGTGACATGCCATTCTGCTGTTACCTGCTTCAATTGTTG	2700
D b	2641	AGACTCTGTTAGGTGACATGCCATTCTGCTGTTACCTGCTTCAATTGTTG	2700
Qy	2701	CAGGATTTGGCTTCTTCATCTCTTAACTGCTATCTCTTCTTATTAATAGTA	2824
D b	2701	CAGGATTTGGCTTCTTCATCTCTTAACTGCTATCTCTTCTTATTAATAGTA	2766
Qy	2761	ATAGTGTGCTGGTCTGTTCTGCTGTTCTGCTGTTCTGCTGTTCTGCTG	2880
D b	2761	ATAGTGTGCTGGTCTGTTCTGCTGTTCTGCTGTTCTGCTGTTCTGCTG	2880
Qy	2821	CACCTTCACATCCAAAGGGTGGTCAAGTCAAGCTGATATAGGTGATCTC	2940
D b	2821	CACCTTCACATCCAAAGGGTGGTCAAGTCAAGCTGATATAGGTGATCTC	2940
Qy	2881	ACGGGTAGTCATAATGAGCCATTGGGAAACTGGAGCTTACGAGACTCTA	3000
D b	2881	ACGGGTAGTCATAATGAGCCATTGGGAAACTGGAGCTTACGAGACTCTA	3000
Qy	2941	ACAGTGGATTGACTCAGCCTGTTGCTGTTGCTGAGACTCTAAGGAGATA	3060
D b	2941	ACAGTGGATTGACTCAGCCTGTTGCTGTTGCTGAGACTCTAAGGAGATA	3060
Qy	3001	GCACATGGTTACATTGGAACTTGTGTTAGGATCATGGCCAGGGTAA	3120
D b	3001	GCACATGGTTACATTGGAACTTGTGTTAGGATCATGGCCAGGGTAA	3120
Qy	3121	ATATGTTTCTTCCTTAAAGATAAAACACAGCTCTACATCCG	3180
D b	3121	ATATGTTTCTTCCTTAAAGATAAAACACAGCTCTACATCCG	3180
Qy	3061	ACTGGACTTTAGGTTAACTAAAGCATGGCTTATCATTTCTGCTGATTG	3240
D b	3061	ACTGGACTTTAGGTTAACTAAAGCATGGCTTATCATTTCTGCTGATTG	3240
Qy	3181	ATGTTACCTCTCTGTTGTTCACTTGTGAACTTCTGTTGAACTAT	3300
D b	3181	ATGTTACCTCTGTTGTTCACTTGTGAACTTCTGTTGAACTAT	3300
Qy	3241	TTGAGGTTAAATGGCTTAACTGAAAGCTTAACTGAAAGCTTAACTGAA	3360
D b	3241	TTGAGGTTAAATGGCTTAACTGAAAGCTTAACTGAAACTTAACTGAA	3360
Qy	3301	AAATATATACTGAAAAATCCACCTTCATCCATCTGGAAATATAGGAA	3420
D b	3301	AAATATATACTGAAAAATCCACCTTCATCCATCTGGAAATATAGGAA	3420
Qy	3361	ATAGTTGAGTTGAGGTGAACTGAAAGTAACTGAACTGTTGAACTAT	3480
D b	3361	ATAGTTGAGTTGAGGTGAACTGAAAGTAACTGAACTGTTGAACTAT	3480
Qy	3421	GAAATAACATTGAACTATTTACTGAAATTATTCAGTAACTTCTCTT	3540
D b	3421	GAAATAACATTGAACTATTTACTGAAATTATTCAGTAACTTCTCTT	3540
Qy	3481	AGCAACCAACATACTCTTAAAGGCTCTAACTTAACTGAACTTCTT	3600
D b	3481	AGCAACCAACATACTCTTAAAGGCTCTAACTTAACTGAACTTCTT	3600
Qy	3541	GACATTACCTGAAAGAAAAATTCTGAGCATTGTTACACCTTATTGAGG	3660
D b	3541	GACATTACCTGAAAGAAAAATTCTGAGCATTGTTACACCTTATTGAGG	3660
Qy	3601	TCCGTTGGATGGCTGGGGCAAGGGTGGTACCTGAAATTCTCTT	3720
D b	3601	TCCGTTGGATGGCTGGGGCAAGGGTGGTACCTGAAATTCTCTT	3720
Qy	3661	AAAACAGAAAGAAAAATTCTGAGCATTGCTTAAAGTGTAACTCTT	3780
D b	3661	AAAACAGAAAGAAAAATTCTGAGCATTGCTTAAAGTGTAACTCTT	3780
Qy	3721	TTGTTAGGGCCATTGAGCTTCTGAGCATTGCTTAACTGAACTTCTT	3780
D b	3721	TTGTTAGGGCCATTGAGCTTCTGAGCATTGCTTAACTGAACTTCTT	3780
Qy	3781	GCAGATTACTGAAATAATATACTTAACTTAACTTAACTTAACTTAA	3780

Db	3781	GCAGATTACTGATAATATACATTATTTAAATATTGGGAGTACTATAATTGTGAGA	3 840	Db	4 861	TTCCTAGAAAAAGATGTTGTTCTATGTCACATAGATGATAAAATAAGTATTG	4 920
Qy	3841	AATGTAATTCGATAATATGTAATATGTAATATATAACACAC	3 900	Qy	4 921	CCAATGCTGTTCACTCTCTAGTGCCAGATCATATTCTGGAAAATTCCTGGTGTGCG	4 980
Db	3841	AATGTAATTCGATAATATGTAATATGTAATATATAACACAC	3 900	Db	4 921	CCAATGCTGTTCACTCTCTAGTGCCAGATCATATTCTGGAAAATTCCTGGTGTGCG	4 980
Qy	3901	ACACATCGAACATCOCGCCACTCTAGATACTAACAGTTGGCTCTGAGGCTTGG	3 960	Qy	4 981	CTTAGCTTGGTTAAAAAAAAAAAAGGATTAACATTAATAAAAG	5 040
Db	3901	ACACATCGAACATCOCGCCACTCTAGATACTAACAGTTGGCTCTGAGGCTTGG	3 960	Db	4 981	CTTAGCTTGGTTAAAAAAAAAAAAGGATTAACATTAATAAAAG	5 040
Db	3901	CTCCGTATAATTATCGCTCTATATTTTAACTCTGAGCNGTAGTTTATAC	4 020	Qy	5 041	TAGTTGAAATTGGGCTCACACAGATATTGAACTTCATCAGTTGTCACATCGAT	5 100
Qy	3961	CTCCGTATAATTATCGCTCTATATTTTAACTCTGAGCNGTAGTTTATAC	4 020	Db	5 041	TAGTTGAAATTGGGCTCACACAGATATTGAACTTCATCAGTTGTCACATCGAT	5 100
Db	3961	CTCCGTATAATTATCGCTCTATATTTTAACTCTGAGCNGTAGTTTATAC	4 020	Qy	5 101	TATGTAACAGTTAGTACCTGATGTTAACAAACCGGAAGTTGAGTGTGAGATCT	5 160
Qy	4021	TTATGTTAAATTATTATTGAAATTACATTATAAAAGTGTGTOCAAGGC	4 080	Db	5 101	TATGTAACAGTTAGTACCTGATGTTAACAAACCGGAAGTTGAGTGTGAGATCT	5 160
Db	4021	TTATGTTAAATTATTATTGAAATTACATTATAAAAGTGTGTOCAAGGC	4 080	Qy	5 161	GCRATGCTATCCATGATGTTAACAAAGTACCTTATATTGACTCTGTTAAACT	5 220
Qy	4081	ATTAATTAATATATGTTAAAGGAAGTACATTTTAAACTGTCCTAGC	4 140	Db	5 161	GCRATGCTATCCATGATGTTAACAAAGTACCTTATATTGACTCTGTTAAACT	5 220
Db	4081	ATTAATTAATATGTTAAAGGAAGTACATTTTAAACTTCAACTGTCCTAGC	4 140	Qy	5 221	TCACTTTAAATCCATTACCAAGTTGAGTTTAAACCTTCATTCCTGGCTGATT	5 280
Qy	4141	TTTGATTTAGGATAATTTCCTGAAGTAGTGGCTTTCATTACTGCTT	4 200	Db	5 221	TCACTTTAAATCCATTACCAAGTTGAGTTTAAACCTTCATTCCTGGCTGATT	5 280
Db	4141	TTTGATTTAGGATAATTTCCTGAAGTAGTGGCTTTCATTACTGCTT	4 200	Qy	5 281	ATGCATCACTCTGTTGGAACCTTTTATTTCATTAGTTTCAAGCTTCTGATT	5 340
Qy	4201	CCCTTGTCTGAAACAGATGCTTACCTTAACTCCTTGTGCTCTGCTCATTACTGCTT	4 260	Db	5 281	ATGCATCACTCTGTTGCTGTGCTTAACTTTTATTTCATTAGTTTCAAGCTTCTGATT	5 340
Db	4201	CCTTAGGTTCTGAAACAGATGCTTACCTTAACTCCTTGTGCTCTGCTCATTACTGCTT	4 260	Qy	5 341	TTTCCTTATTGGCTGCTGTGCTTATTTCATTAGCTTAACTAGTGTATA	5 400
Qy	4261	ATATGCTTAAATCACCATATAAGGAAAAAATTGGTACAGAACAAATAAAC	4 320	Db	5 341	TTTCCTTATTGGTCTGTGCTTATTTCATTAGCTTAACTAGTGTATA	5 400
Db	4261	ATATGCTTAAATCACCATATAAGGAAAAAATTGGTACAGAACAAATAAAC	4 320	Qy	5 401	TTGGTTAAATTGGACAGTGGCTTTTAAACCAATACACTGTAATGAGCTAGCA	5 460
Qy	4321	CTATTTTAAATGAAATCACAATAGCAAGTGGCTTAACACTTATTCTGTTAA	4 380	Db	5 401	TTGGTTAAATTGGACAGTGGCTTTAAACCAATACACTGTAATGAGCTAGCA	5 460
Db	4321	CTATTTTAAATGAAATCACAATAGCAAGTGGCTTAACACTTATTCTGTTAA	4 380	Qy	5 461	GAGCGATTGCTTATAAGCTTAACTAACTGTTAAACGATCTGAACTAGCA	5 520
Qy	4381	ATGACTTAAAGAATGCTCAAAATTAGTGAAGTGGCTTAACACTTATTCTGTTAA	4 440	Db	5 461	GAGCGATTGCTTATAAGCTTAAAGAGTACATATTTCAGGATCTG	5 520
Db	4381	ATGACTTAAAGAATGCTCAAAATTAGTGAAGTGGCTTAACACTTATTCTGTTAA	4 440	Qy	5 521	TCATCTAGTGTGCAATACATATAAGCTTAACTAACTGTTAAACGATCTG	5 580
Qy	4441	ACTGCTAAATTAAAAAGGGAGTACCAAGATTAACTAAAGGAAATGGG	4 500	Db	5 521	TCATCTAGTGTGCAATACATATAAGCTTAACTAAAGGAAATGGTGTG	5 580
Db	4441	ACTGCTAAATTAAAAAGGGAGTACCAAGATTAAAGGATTAAAGGATTTGGAAATGGG	4 500	Qy	5 581	CAACCTCTATACATACATATAAGCTTAACTAAAGGAAATGGACTCTAACTTAA	5 640
Qy	4501	ATAGGCCATATATGTTATGTTATGTTATGTTATGTTATGTTATGTT	4 560	Db	5 581	CAACCTCTATACATACATATAAGCTTAACTAAAGGAAATGGACTCTAACTTAA	5 640
Db	4501	ATAGGCCATATATGTTATGTTATGTTATGTTATGTTATGTTATGTT	4 560	Qy	5 641	CAGAACCTCTTATAGGTATTGGCTTTTATTTAGACTTATAATCAAGTGTG	5 700
Qy	4561	ACACATACATACATTAATCTGCCCTGCTGAATTAACATGGAGCACATCTCA	4 620	Db	5 641	CAGAACCTCTTATAGGTATTGGCTTTTATTTAGACTTATAATCAAGTGTG	5 700
Db	4561	ACACATACATACATTAATCTGCCCTGCTGAATTAACATGGAGCACATCTCA	4 620	Qy	5 701	TAGATGATATATGTTAGGCTTTGATATAATAGCTTGTGAAATAGGATATG	5 760
Qy	4621	GGGCACAGCTGTTAAATTGGAGCTTAACTTCACTGTTACCTCGCTGTT	4 680	Db	5 701	TAGATGATATATGTTAGGCTTTGATATAATAGCTTGTGAAATAGGATATG	5 760
Db	4621	GGGCACAGCTGTTAAATTGGAGCTTAACTTCACTGTTACCTCGCTGTT	4 680	Qy	5 761	TATTTAAACAGGTACATTTCATAGTTTACAATGGTTTATTTACAAGGTTACGTG	5 820
Qy	4681	CCACCCCCAGACTTGAATAACACTCGACTAACTTCACTGTTACCTCGCTGTT	4 740	Db	5 761	CAGGTACATTTCATAGTTTACAATGGTTTATTTACAAGGTTACGTG	5 820
Db	4681	CCACCCCCAGACTTGAATAACACTCGACTAACTTCACTGTTACCTCGCTGTT	4 740	Qy	5 881	TCAAGGACACTGCTTCCTCCAGTACCTAAGTGTACCTAAGTGTACCT	5 940
Qy	4741	AAATTGCTGTTGAGTCCATGTCACATGGGAAATTCACTTAAAGGATTTG	4 800	Db	5 881	TCAAGGACACTGCTTCCTCCAGTACCTAAGTGTACCTAAGTGTACCT	5 940
Db	4741	AAATTGCTGTTGAGTCCATGTCACATGGGAAATTCACTTAAAGGATTTG	4 800	Qy	5 941	CTAGCCTACTGAGATCCATGTCACATGGGAAATTCACTTAAAGGATTTG	6 000
Qy	4801	ACTCGCTTATGTTAACTGCAATGTTCACTGTTACCTAAGGAAATTG	4 920	Db	5 941	CTAGCCTACTGAGATCCATGTCACATGGGAAATTCACTTAAAGGATTTG	6 000
Qy	4861	TTCCCTAGAAAAAGATTTGTTCTATGTCACATGTCACATGGGAAATTG	4 920				

QY	6001	TTAGCAATGGCTGTAGTCTGTAATGGCTAAGTGTGAATTGGCTT	6060	QY	7081	GTCGTATAAGCTCTGTAAAATGAAATCACATTGAAAGCTGGAGTGACTACATT	7140		
Db	6001	TTAGCAATGGCTGTAGTCTGTAATGGCTAAGTGTGAATTGGCTT	6060	Db	7081	GTCGTATAAGCTCTGTAAAATGAAATCACATTGAAAGCTGGAGTGACTACATT	7140		
QY	6061	TTCACACTGTGTOAGGCCATGCTAACACAGATAAGTGTGATCACATG	6120	QY	7141	AAACATTCACATCCAAAGAGCGAGCATATTATGCACTATGAGATTAAATATA	7200		
Db	6061	TTCACACTGTGTOAGGCCATGCTAACACAGATAAGTGTGATCACATG	6120	Db	7141	AAACATTCACATCCAAAGAGCGAGCATATTATGCACTATGAGATTAAATATA	7200		
QY	6121	ATCTACATAGTTGCACTTGTGCAAAAACCCATAGGCAAGCTCTTGGATCATA	6180	QY	7201	ATTTGCTGTAATTAATTAATAGTACTCAGCTTCTGGCTTACAGTTATGTTGCG	7260		
Db	6121	ATCTACATAGTTGCACTTGTGCAAAAACCCATAGGCAAGCTCTTGGATCATA	6180	Db	7201	ATTTGCTGCTTAATTAATTAATAGTACTCAGCTTCTGGCTTACAGTTATGTTGCG	7260		
QY	6181	GAAGAAGATCATGCACACAGATTTGTAAGGCAACTCACATTGCACTAGGGCTT	6240	QY	7261	TGTAAGATAAGATAAGTGTGAATTCCACAAATAATGAAATAATCTGTGCC	7313		
Db	6181	GAAGAAGATCATGCACACAGATTTGTAAGGCAACTCACATTGCACTAGGGCTT	6240	Db	7261	TGTAAGATAAGATAAGTGTGAATTCCACAAATAATGAAATAATCTGTGCC	7313		
QY	6241	TCTATGATGTTGTCACCCCTCTGAGGATGAAAGGCACTGCTCTAGGGCT	6300	RESULT 2					
Db	6241	TCTATGATGTTGTCACCCCTCTGAGGATGAAAGGCACTGCTCTAGGGCT	6300	AAV18481	ID	AAV18481 standard; cDNA; 2334 BP.			
QY	6301	AGAAATGACACAACATTTGCTTAATGCTTACATTAAGCTTCAAGTGGT	6360	AAV18481;	XX				
Db	6301	AGAAATGACACAACATTTGCTTAATGCTTACATTAAGCTTCAAGTGGT	6360	AC	XX				
QY	6361	ACAGATGACCAAGATGAAATGATGTTATCTCTAGAAACACTCTTCAATATTGGA	6420	DT	18-AUG-1998	(first entry)			
Db	6361	ACAGATGACCAAGATGAAATGATGTTATCTCTAGAAACACTCTTCAATATTGGA	6420	XX					
QY	6421	TCTATGCTGCTATGTACTGGCTCAAACTCTCTCTCTCTCTCTCTCTCT	6480	XX					
Db	6421	TCTATGCTGCTATGTACTGGCTCAAACTCTCTCTCTCTCTCTCTCTCT	6480	XX					
QY	6481	ATTCACTGCTGTTTATTCATGAAAGGACTACATTTAGGTCTAAAGTGTACAA	6540	OS					
Db	6481	ATTCACTGCTGTTTATTCATGAAAGGACTACATTTAGGTCTAAAGTGTACAA	6540	Homo sapiens.					
QY	6541	ATATTTTATACTGTGACTTAATTTGCTTAACAACTTTACACCAATGTATT	6600	XX					
Db	6541	ATATTTTATACTGTGACTTAATTTGCTTAACAACTTTACACCAATGTATT	6600	XX					
QY	6601	CATGTGCACTGCAAAGGAGATCTGGACATGAAATGTTACAGAACACCAAGCTT	6660	Key					
Db	6601	CATGTGCACTGCAAAGGAGATCTGGACATGAAATGTTACAGAACACCAAGCTT	6660	CDS					
QY	6661	TGTCCCAAGCTGACTCAAGATGAAAGTGGCTTATTAATGGCTGGAGT	6720	Location/Qualifiers					
Db	6661	TGTCCCAAGCTGACTCAAGATGAAAGTGGCTTATTAATGGCTGGAGT	6720	803 .2194					
QY	6721	GAAGAACATGCTGTACTAAAGCCCTTGACATCTGAACTAGCTGAAATCCATT	6780	FT					
Db	6721	GAAGAACATGCTGTACTAAAGCCCTTGACATCTGAACTAGCTGAAATCCATT	6780	FT					
QY	6781	AGGAAACGGATTGCTCATCTGAACTAGCTGAAATGGTATTAGCTGAAATTCATT	6840	FT					
Db	6781	AGGAAACGGATTGCTCATCTGAACTAGCTGAAATGGTATTAGCTGAAATTCATT	6840	FT					
QY	6841	ATGAAAAAGGTGCTGAGTACTCTAAAGCTTCAATGATAAACTTTGAGTCAA	6900	PR					
Db	6841	ATGAAAAAGGTGCTGAGTACTCTAAAGCTTCAATGATAAACTTTGAGTCAA	6900	PR					
QY	6901	AATAGAAAAGAAAAAATCTGCACTCCAGCCGAATTGTTATTTGATTTAA	6960	XX					
Db	6901	AATAGAAAAGAAAAAATCTGCACTCCAGCCGAATTGTTATTTGATTTAA	6960	CC					
QY	6961	AATTGCTTATCTGTAATATGGAAATCAGTGGCTTATCATGTGTTACTTAA	7020	CC					
Db	6961	AATTGCTTATCTGTAATATGGAAATCAGTGGCTTATCATGTGTTACTTAA	7020	CC					
QY	7021	ATGATTCACAAACTACTGTGTTGATTAATAGACAAAGATCATATTGTTG	7080	CC					
Db	7021	ATGATTCACAAACTACTGTGTTGATTAATAGACAAAGATCATATTGTTG	7080	CC					

Claim 17: Pages 82-85; 118PP; English.

The present sequence represents the human zinc finger protein (hZAC) cDNA isolated from the human pituitary gland cDNA library using BOP1 cDNA (AAW18480) as a probe. Therefore the hZAC protein is the human homologue of BOP1 (AAW47761). The hZAC protein displays a tumour suppressing activity when it was constitutively and inducibly expressed in tumour cells. The hZAC cDNA and the protein it encodes are claimed to be useful in the preparation of therapeutic compositions, useful for treating, preventing or delaying the recurrence of a tumour or neuronal disorders, e.g. genetic diseases or acquired degenerative encephalopathies such as Alzheimer's disease. The hZAC protein is also claimed to be able to induce apoptosis resulting in inhibition of tumour cell growth, to suppress tumour formation, to induce G1 arrest of the cell cycle and to act as nuclear transcription factor.

XX	SQ Sequence 2334 BP; 569 A; 636 C; 550 G; 579 T; 0 other;	PD 16-MAY-2002.
SQ	Query Match Score 351.6; DB 19; Length 2334;	XX 08-NOV-2001; 2001WO-US46816.
Best Local Similarity 4.8%; Pred. No. 6e-57;	XX 09-NOV-2000; 2000US-247457P.	
Matches 490; Conservative 0; Mismatches 209; Indels 3; Gaps 1;	XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.	
Qy 543 GAAACGTGTAAGCTGGTGAACAAACCAAGAAAACACTTCCCTGCCAACACTGTGTGACAA 602	XX PA PT -Lee RT;	
Db 775 GAAAAGGAGGTGTGAGAGCAAAGCCATGGCACCTTCCTGCCAGTGTGTGCAA 834	XX XX WPI; 2002-590446/63.	
Qy 603 GGCTTTPAAGTGTGAAATTAAAGTTCACTCTCAGAGGAGAGGCC 662	XX DR Diagnosing cardiovascular conditions e.g. stroke, myocardial infarction	
Db 835 GACGTTCTCACCTGAGAAGTCAAGATTCACATTATGCCACTCGAGGCC 894	PT in a subject, by detecting expression of a nucleic acid molecule such	
Qy 663 CTACAAGTGCATACACAGACTGCAACAGGCTTCTTAAGTACAATTACAAG 722	PT as Fit-1, AA892598 and Mrg-1 in biological sample from	
Db 895 GTACAAGTGTGCGTACTGACTGTGGAAAGCTTCCAGATAATTGTGATG 954	PT subject -	
Qy 723 GCACATGGTCACTCATTCCTCTGAAACACCAAAAGTGTAAATTCTGAGAAATT 782	XX Disclosure; Page 107-108; 113pp; English.	
Db 955 GCATATGGTACCCATTCCCAGAAATCTCAGGTTGCTACTTGAGAGCCTT 1014	XX The invention relates to diagnosing a cardiovascular condition	
Qy 783 TCACCGGAAAGATGATCTCAAGATCACCTCCATACAGACGCCCTAACAAAGAGACCTT 842	CC characterised by aberrant expression of a nucleic acid molecule (I) such	
Db 1015 CAACCGGAAAGACCACTGAAACCCATCCAGAACCCACAAATAGGCTT 1074	CC as Fit-1, CD44, Lot-1, AA892598 and Mrg-1 or its expression product. One	
Qy 843 TAAGTGCAGAAATGTGGAGAACTCATACAACTGGATTAAACGTACTGGC 902	CC method involves contacting a biological sample from a subject with an	
Db 1075 TGGGTGTAGGAGTGTGGAGAGGTACACCATGTGGGTATAAGGGACCTGGC 1134	CC agent which specifically binds to (I), its expression product or a	
Qy 903 CCTGCATGGCCACAAAGTGTGACCTTACCTGTAAACTTTGAAAG 962	CC fragment of an expression product and measuring the amount of bound	
Db 1135 CCTCCATGGCCACAGTGGGACTCTGGGACTCTGGCTCTGGTAGGAG 1194	CC agent. The method is useful for diagnosing a cardiovascular condition	
Qy 963 CACGGAGTGTCTTGGAGCACCTTAATCTCATGC -- AGGCAAGTGTCTGGTGGCT 1019	CC such as myocardial infarction, stroke, arteriosclerosis, heart failure,	
Db 1195 CACCGAGGTCTACGGACCCATCGAACGACCTCAAAAGCCATCGAAC 1254	CC and cardiac hypertrophy. Methods useful for determining the stage of	
Qy 1020 TAAAGAAAAAAAGCACATGTGCAACATGTGATCGCGTCTAACCGAAGGACT 1079	CC cardiovascular condition and pharmaceutical compositions for treating the	
Db 1255 CAAGGAAAGAGGACCACTGGGACCACTGGGACCACTGTGAAAGTGT 1314	CC above cardiovascular conditions are also provided. The present sequence	
Qy 1080 CGCGGAGACATGGTGTGCAACTCGGAAGAAGGACTTCTCTCTCAATTTGACA 1139	CC represents a human ZAC zinger finger protein cDNA.	
Db 1315 GCGACCCACCTGGGCTCACAGATGAAAGACTCTGTGCCCCATGCTGCCA 1374	XX Sequence 2828 BP; 701 A; 722 C; 651 G; 754 T; 0 other;	
Qy 1140 GAGATTGGCGAACAGGATCACCTGACATATGAAAGAGTCAACATCAAGACT 1199	Query Match Score 351.6; DB 24;	
Db 1375 GAGATTGGGCAACAGATCACCTCACCGGCATACAAAGAACCTACAGGACT 1434	Best Local Similarity 65.8%; Pred. No. 6.3e-57;	
Qy 1200 TCTGAAGGTGCTAAACAGAACAGTGGATTCTGACCCATT 1241	Mismatches 0; Conservatve 490; Gaps 3; Gaps 1;	
Db 1435 GATGAAAGAGGCTGTGAGACGGAGACCTCTGACACCTT 1476	Qy 543 GAAAGTAAAGCTGTGTTGAAACCCAAACCTGTGACAACTTGTGACAA 602	
Db ABN86525; RESULT 3	Db 1269 GAAAGGAGGTGTGAGAGCAAAGCCATGGCCAGTTCAGTGTGTTGGAA 1328	
XX ABN86525 standard; cDNA; 2828 BP.	Qy 603 GGCCTTAAACAGTGTGAAATTAAGGTTCACTCTAACCTCACAGGAGGAGGCC 662	
XX ABN86525;	Db 1303 GGCCTTCTAACCTGAGAGTTCACGTTCAATTTCCACTCCAGGGACGGCC 1398	
AC	Qy 663 CTACAAAGTGCATACAAAGAACGACTGCAACAGGCTTCTGAGAAATTAAGTACAATGAAAG 722	
XX ABN86525;	Db 1389 GTACAAAGTGTGTGCGCTGACTGTGCAAAAGCTTGTGTTCCACATATAATGATGAG 1448	
DT 21-OCT-2002 (first entry)	Qy 723 GCACATGGTCACTCATTCCTCTGAGAAACCCCAAGTGTGAAATATTGTGAGAAATGTG 782	
XX Nucleotide sequence of human ZAC zinger finger protein cDNA.	Db 1449 GCATATGGTCACTCCCGAAGATCTCCCAAGTGTGCTACTGTGAAAGACGTT 1508	
DE	Qy 783 TCACCGGAAGATCATCTGAGAACTACCTCCATACACAGGCTAACAAAGGAGACGTT 842	
XX Cardiovacular; Fit-1; CD44; Lot-1; AA892598; Mrg-1; cardiac; human;	Db 1509 CAACCGGAAGAACGACCTGAAACCCACCTCCAGCCAGCACCCTAGCCCT 1568	
KW cerebroprotective; antiarteriosclerotic; ZAC; zinger finger protein; ss.	Qy 843 TAAGTGCAGAAAGTGTGGCAAGAACATCAAACTGACGCTTAACTTGAAAG 902	
XX Homo sapiens.	Db 1629 CCTCCATGGCCACAGTGGGACTCTGGGCTCTGGGCTCTGGGCTCTGGG 1688	
OS WO200238794-A2.	Qy 963 CACGGAGGTGTCTGGGACCTTAATCTCATGC - AGGCAAGTGTGTGGGGT 1019	
XX	Db 1689 CACCGAGGTGTCTGGGACCTCAAGCCATGGAAAGAACGCCCTAGCGAAC 1748	
DN	Qy 1020 TAAAGAAAAAGCAGGACTGTGGAAATTTGATCGCGGTTCAACCCGAAAGGTTG 1079	
XX	Db 1749 CACGGAGGTGTCTGGGACCTCAAGCCATGGAAAGAACGCCCTAGCGAAC 1808	

Query Match Score 294.4; DB 19; Length 2790;
 Best Local Similarity 64.0%; Pred. No. 4.2e-46;
 Matches 455; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy 557 GTGAAACCAAAACCAAGAAAACACTTCTTGCCAACTGTGACACGGCCTTAACAGTG 616
 Db 528 GAGAACCAAAGCCATGGCTCCATTGGCTGTCAAATGGCAAGTCCTTGTCGACCC 587

Qy 617 TTGAGAAATTAAAGGTCACTCTTCACTACAGTGTGATAC 676
 Db 588 TTGAGAAGTTACCATTCATTTCCACTCCAGGCCATTAAGTGTGCGA 647

Qy 677 AACAGAGTGCACCAAGGCCACATGGTACTC 736
 Db 648 AGCCTGACTGTGCGCAAAAGCTTGTGTCMCAACTATAGCTGATGAGACATGGCAC 707

Qy 737 ATTCCTCCGAGAAACCAACGTCAGTGTAAATTATTTGAGAAAATGTTAACCGGAANGATC 796
 Db 708 ACTCGCCACAGAGATTCACTGAAAGCATTCACGTGAAAGCATTCACCCGAAGAACCC 767

Qy 797 ATCTGAAACATAACCTCCATAACAGCCCTAACAGAGGTGTTAACAGAGGTGAGAGAT 856
 Db 768 ACCTGAAACACCTCGAACCCACGTTCCACAGATCTCCAGCGTGTGACATT 827

Qy 857 GTGGCAAGAAACTACAATAACCAAGTGGATTAAACCTCACTTGGCCTTGCAATGCCGAA 916
 Db 828 GCGGCAAGAAAGTACACATGGCATTGCTGGCTGCTGACGTGGCGA 887

Qy 917 CAACTGGTGAACCTCACCGCTAAAGTATGTTGAAACCTTTGAAAGCACGGGAGTGTCTC 976
 Db 888 GCATGGCAGATCCTCACCTGTGGCTGGCTGGACGGAGACCGAGTCCTGTC 947

Qy 977 TGGAGACCTTAATCTCATGC -- AGGGAAGTGTGTGTTGGGGTTAAGGAAAGG 1033
 Db 948 TGGACCACTCAAGTCTACGGCGAAAGAAAAGGCCAACCGGACCCCAGGAGAAGAAT 1007

Qy 1034 ACCAGTGGCAGACATTGTGATCGGGTCTACACCGAAAGGATGTGCTGGCACCCTG 1093
 Db 1008 ACCAGTGGCAGACATTGTGATCGGGTCTACACCGAAAGGATGTGCTGGCACCCTG 1067

Qy 1094 TGGGCCACACTGGAAAGGAAGGACCTCTCTGTCAGTAGTGTGCAAGAGATTGGCGAA 1153
 Db 1068 TGGTCCCAACAGGTGCAAGGCTCTGTCAGTTGTGCAAGATTTGGGCA 1127

Qy 1154 AGGATCACCTGGACTGCAATATGAGAAAGTCAATCAGACTCTCAGAGTCAAA 1213
 Db 1128 AAGAACCCCTCACTGGTCAACCAGAACCCACTCCAGGAGETGTGATGCAAGATA 1187

Qy 1214 CAGAACCATGGATTCTCTGACCCATTACCTGCAATGTGTCT 1257
 Db 1188 TGAGGGAGGATTACAGGCAATTCCAACTTGCCCT 1231

RESULT 8
 ID ABN86526 standard; cDNA; 3975 BP.
 XX ABN86526;
 XX DT 21-OCT-2002 (first entry)
 XX DR Nucleotide sequence of mouse ZAC zinger finger protein cDNA.
 XX Cardiovacular; Fit-1; CD44; Lot-1; AA82598; Mrg-1; cardiac; mouse;
 XX cerebroprotective; antiartherosclerotic; ZAC; zinger finger protein; ss.
 OS Mus musculus.
 XX WO20038794-A2.
 XX PD 16-MAY-2002.
 XX PP 08-NOV-2001; 2001WO-US46816.

XX PR 09-NOV-2000; 2000US-247457P.
 XX PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX PI Lee RT;
 XX DR WPI; 2002-590446/63.
 XX PT Diagnosing cardiovascular conditions e.g. stroke, myocardial infarction
 PT in a subject, by detecting expression of a nucleic acid molecule such
 PT as Fit-1, CD44, Lot-1, AA82598 and Mrg-1 in biological sample from
 PT subject -
 XX Disclosure; Page 108-109; 113pp; English.
 PS
 XX The invention relates to diagnosing a cardiovascular condition
 CC characterised by aberrant expression of a nucleic acid molecule (I) such
 CC as Fit-1, CD44, Lot-1, AA82598 and Mrg-1 or its expression product. One
 CC method involves contacting a biological sample from a subject with an
 CC agent which specifically binds to (I), its expression product or a
 CC fragment of an expression product and measuring the amount of bound
 CC agent. The method is useful for diagnosing cardiovascular condition
 CC such as myocardial infarction, stroke, arteriosclerosis, heart failure,
 CC and cardiac hypertrophy. Methods useful for determining the stage of
 CC cardiovascular condition and pharmaceutical compositions for treating the
 CC cardiovascular conditions are also provided. The present sequence
 CC represents a mouse ZAC zinger finger protein cDNA.
 XX Sequence 3975 BP; 1010 A; 1001 C; 880 G; 1084 T; 0 other;
 SQ Score 4.0%; Score 294.4; DB 24; Length 3975;
 Best Local Similarity 64.6%; Pred. No. 4.5e-46;
 Matches 455; Conservative 0; Mismatches 246; Indels 3; Gaps 1;
 Qy 557 GTGAAACCAAAAGAGAAAAAAACTTCTCTGCAACTGTGCAAGGGCCTTAACAGTG 616
 Db 235 GAGAACCAAGGCATGGCTCCATTCACCTCCAGTGGCTGTCBAAAATGTGGAAAGTGCCTCACC 294
 Qy 617 TTGGAAATAAAGGTTCACTCTCACTCTCAGCTACATGGCTGAGAGGAGGCTAACATGGCTACTC 676
 Db 295 TGGGAAAGTTCACATTCAAAATTCACCTCCAGTGGCTGTCBAAAATGTGGAAAGTGCCTCACC 294
 Qy 677 AACAGACTGACCAAGGCTCTTCTAAAGTACAATAAACAGGCAATGGCTACTC 736
 Db 355 AGGCTGAGTGTGGAAAAGCCTCTCCAAAGTAAAGTGTGATAAGAACATGGCTACAC 414
 Qy 737 ATTCTCCTGAGAAACCCAACAAGTAAATTATGTGAGAAATGTTCACCGGAAAGATC 796
 Db 415 ACTCCCAZGAAATTCAACGAGTCACTGTGACCTGACATCTGAGAAAGACTTCACCTGAGAA 474
 Qy 797 ATCTGAGAACTCACCTCCATACAGGCTTCAAGGCTTAAGAGGAGCTTAAGTGAGAAAT 856
 Db 475 ACCTGAGAACCCCTCAACCCAGATECCAAAGATCTCTACGGGTGTGAGCATT 534
 Qy 857 GTGGCAAGAACTACAATAACCAAGCTTGGATTAAAGTCACTTGCCCTTCGATGCCGCAA 916
 Db 535 GCGGCAAGAACTACCAACCATGGGGTACAAAGGGACCTGACCTGGCTACTGGCGA 594
 Qy 917 CAAGTGTGAGCTCACCTGTAAGGTGTTGCAAACTTGTGAAAGCACGGGAGTGCTTC 976
 Db 595 GCAATGGCGATCTCACCTGGGGTGTGACCCCTGAGCTGGCTGACCCCTGAGGAA 654
 Qy 977 TGGAGACCTTAATCTCATGC - AGGAGTGTGCTGGGGTTAAGGAAAGGATGTCGGCTG 1033
 Db 655 TGAACACCTCAAGTCTCACSGGAGAAAGGCCAACCCAGGCAACCCAGGAA 714
 Qy 1034 ACCAGTGGCAACATTTGATGCCGTTCTACACCCGAAGGATGTCGGGAGAATGG 1093
 Db 715 ACCAGTCGACCACTGGAAAGGACTCTCTGCACTATTGTGATGATGCTACCCGGACCTGG 774
 Qy 1094 TGGTGCACACTGGAAAGGAAGGACCTCTCTGTCAGTAGTGTGCAAGAGATTGGCGAA 1153

Db	775	TGGTCCACACAGGATCCAAGGACTCCTGTCTAGTTCTGCCAGAATTGGCGCA	834	Query Match 3 ; Score 277.6; DB 19;
QY	1154	AGGATCAGCTGACTGACATGACATATGAAAGAGTCACATCAAAGGCTTCAAA	1213	Best Local Similarity 71.9%; Pred. No. 6.2e-43;
Db	835	AAGGACCATCTCACTGTCACACCAGAAAGCCACTCCAGGAGCTGATGCAAGATA	894	Matches 377; Conservative 0; Mismatches 144; Indels 3; Gaps 1;
QY	1214	CAGAACCACTGGTGAATTCCCTTGACCCATTACCTGCAATGCTCT 1257		QY 721 AGGGCATGCTACTCATCTCCCTAGAAAACCCAAAGTGTAAATATGAGAAAATG 780
Db	895	TGCGGGAGGAGATACCGAGAACTTCCAACTCATGGCCT	938	Db 171 AGGGATATGCTTACCATCTCCCAAGAAATCTTACCTGTCACGTGAGACG 230
RESULT 9				Qy 781 TTTCACCGGAAAGATCATCTGAAGATCACCTCTTACACAGCCCTAACAAAGAGACG 840
ID AA129269	standard;	cdna;	2561 bp.	Db 231 TTCAACCGGAAAGGACCTGAATAACCTCAGACCCAGCCCAACAAATGGCC 290
XX				Qy 841 TTAAAGTGGGAAGATGTGGCAGAACTACAATAACAGCTTGAATTAAACGTCACTTG 900
AC				Db 291 TTTGGTGTGAGGTGTGGAAAGAACATGACATGAACTGCTGCTATAAGGGCACCTG 350
DT	21-AUG-1998	(first entry)		Qy 901 GCCTTGCAATGCCAACACAGTGGTACCTCACCTGTAAGGTATGTTCGAAACTTTTGAA 960
XX				Db 351 GCCCCATGGCCACGGTAGGGACCTCACCTGACACCTAACGGCTAGCTGGCTGAGCTAGGG 470
DE	Nucleotide sequence of human PLAG2.			Qy 961 AGGACGGAGAGTGTCTTGAGAACCTTAATCTCATGC--AGGCAAGTCGCTGCGGG 1017
XX				Db 411 AGCACCGAGGTGTACCTGACACCTAACGGCTAGCTGGCTGAGAGGAGGCCCCTAGGCCA 470
KW	Human PLAG2 gene; PLAG2; tumourigenesis gene; T-gene; PLAG1; CTNNB1; cancer; inhibition; ss.			Qy 1018 GTAAAGAAAAAAAGACACAGNGGACATSTGATGCCGTTCTACCCGAAAGAT 1077
KW	antibody; benign tumour; malignant tumour; leukaemia; lymphoma;			Db 471 ACCAGGAAAAGAGGACCAAGTGAAGATGCTTACCCGGAAAGAT 530
KW	cancer; inhibition; ss.			Qy 1078 GTCCGGAGACACATGGGTTGACACTCTGGAAAGGACTTCCTCTGTAGTATGTRGCA 1137
XX				Db 531 GTGCGACGCCACCTGTGGTCCACATGGATGCAAGGACTTCCTGTCAGTCTGRC 590
OS	Homo sapiens.			Qy 1138 CAGAGATTGGCGAAAGGATACCTGACTCAGCTGGACACTCTGGAAAGAGTCAATCAGAG 1197
XX				Db 591 CAGAGATTGGCGCAAGGTTACCTCACGGGATACCCACGGCACTCACGGAG 650
Key	Location/Qualifiers			Qy 1198 CTTCTGAGGTCAAACAGAACCTGGATTCTCTGTCAGACCTT 1241
CDS	177..1416			Db 651 CTGATGAAAGAGGCTTGAGACCCAGGACCTTCAGAGACCTT 694
FT	/*tag= a			RESULT 10
FT	/product= "PLAG2 protein"			ABT11028
XX				ID ABT11028 standard; cDNA; 2561 BP.
EP825198-Al.				XX ABT11028;
PN				AC ABT11028;
XX				XX
XX				DT 04-DEC-2002 (first entry)
PD	25-FEB-1998.			XX DE Human breast cancer associated coding sequence SEQ. ID NO: 1162.
XX				XX KW Human; breast specific gene; breast cancer; differential expression; cytosolic; gene therapy; gene; ss.
PT	17-JAN-1997;			XX OS Homo sapiens.
XX				XX PN WO200259271-A2.
PF	22-AUG-1996;			XX PD 01-AUG-2002.
PR	96EP-0202339.			XX PF 25-JAN-2002; 2002WO-US02176.
XX				XX PR 25-JAN-2001; 2001US-263757P.
PA	(KULE-) KU LEUVEN RES & DEV.			XX PR 25-APR-2001; 2001US-266090P.
PA	(UYGO-) UNIV GOETEBORG'S HOLDINGBOLAGET AB.			XX PR 23-MAY-2001; 2001US-292517P.
PA	(LEUV-) LEUVEN RES & DEV.			XX PA (GENE-) GENE LOGIC INC.
XX				XX PI Orr MS, Nation M, Diggans JC, Zeng W;
PI	Kas KP, Sternman KGD, Van De Ven WJM, Voz ML;			XX XX
XX				XX DR WPI; 2002-674803/72.
DR	1998-132252/13.			XX XX
DR	P-PSDB; AAW37349.			XX XX
XX				XX XX
PT	New tumourigenesis T-genes and proteins - useful for, e.g. preparing			XX XX
PT	antibodies for clinically diagnosing cells having non-physiological			XX XX
PT	proliferative capacity such as lipoblastomas			XX XX
XX				XX XX
PS	Claim 4; Fig 8; 71pp; English.			XX XX
XX				XX XX
CC	This is the nucleotide sequence of the human PLAG2 gene. It is a			XX XX
CC	tumourigenesis gene (T-gene), which is isolated in the form of PLAG1,			CC XX
CC	PLAG2, and CTNNB1 genes. Their proteins can be used as a starting			CC XX
CC	point for preparing antibodies for clinically/medically diagnosing			CC XX
CC	cells having a non-physiological proliferative capacity as compared			CC XX
CC	to wild type cells, where the former cells are selected from both			CC XX
CC	benign and malignant tumours, as well as leukaemia and lymphomas.			CC XX
CC	Derivatives of the T-gene are also used in the diagnosis and			CC XX
CC	preparation of therapeutic compositions for the treatment of cancers,			CC XX
CC	such as nucleic acid derivatives, and antibodies. The T-gene may be			CC XX
CC	used as a starting point for designing suitable expression-modulating			CC XX
CC	compounds or techniques for the treatment of non-physiological			CC XX
CC	proliferation phenomena in humans or animals. Expression inhibitors			CC XX
CC	of the T-gene can be used in the treatment of diseases involving			CC XX
CC	benign or malignant tumours.			CC XX
SQ	Sequence 2561 BP; 704 A; 624 C; 516 G; 716 T; 1 other;			PT Diagnosing breast cancer in a patient comprises detecting the level of

gene expression in cell or tissue samples, where a differential gene expression is indicative of breast cancer -

CC Claim 1; SEQ ID NO 1162; 260pp + Sequence Listing; English.

CC The present invention relates to methods of diagnosing breast cancer CC patient, which comprise detecting the level of expression in a tissue CC sample of two or more genes selected from those shown in ABT0867- CC ABT1112, where a differential expression of the genes indicates brea CC cancer. The methods are useful in diagnosing, treating, detecting the CC progression, and in monitoring treatment of breast cancer in patients CC progression. The methods are also useful as a screening tool for agents that modul CC the onset or progression of breast cancer. The breast cancer genes may CC be used as diagnostic markers for the prediction or identification of the CC malignant state of breast tissue, for confirming the type and progress CC of cancer, and for drug screening and assays. The present sequence is CC coding sequence of the invention.

CC Note: The sequence data for this patent did not form part of the prin CC specification, but was obtained in electronic format directly from WI CC at ftp://wipo.int/pub/published_pct_sequences.

Qy 1190 ATCGAGACTCTGAAAGTCAAAACACAGTGGATTCCCTGACCCATT 1241
 Db 804 CACAGGACTGTGAAACAGAGCCTTGCAGACCGAACCTTGACCTT 855

RESULT 12
ABA52644/C
ID ABA52644 standard; DNA; 475 BP.

XX Human foetal liver single exon nucleic acid probe #949.
 DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX Human; foetal liver; gene expression analysis in human heart cell sample.
 DE Human; gene expression; heart; microarray; vascular system; probe;
 XX Homo sapiens.
 OS WO200157277-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00669.
 XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-020456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000US-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-483447/52.
 XX PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX PS Claim 1; SEQ ID NO 949; 639pp + sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 SQ Sequence 475 BP; 75 A; 129 C; 142 G; 129 T; 0 other;

Query Match 3.7%; Score 267; DB 22; Length 475;
 Best Local Similarity 74.1%; Pred. No. 4.3e-41;
 Matches 352; Conservative 0; N mismatches 120; Indels 3; Gaps 1;

Qy 733 ACTCATTCCTGAGAAACACAGTGTAAATTATGTTGAGAAAATGTTTCACCGGAA 792
 Db 475 ACCCATCTGAAATCACCTCCATACAGAACCTTAAGAGAGCTTAAGTGGCA 416

Qy 793 GATCATCTGAGAAATCACCTCCATACAGAACCTTAAGAGAGCTTAAGTGGCA 852
 Db 415 GACCACTGAAAACACCTCCAGACCCAGACCTCAAAATGGCCTTGGGTGAG 356

Qy 853 GAATGTCGGAGAGACTACATACAAAGTTAAAGTCGACTTGCCTGCATGCC 912
 Db 355 GAGTCGGAGAGACTACATGCTGGCTATAAGGCACCTGCCTCCATCG 296

Qy 913 GCAAACAAGTGGTGAACCTCACCCTGAGCTTGTGAAACTTGTGAAAGCAGGGACTG 972
 Db 295 GCGAGCAGTGGGACCTCACCTGCGGTCTGCCCCCTGGAGCTAGGACCCAGGTG 236

Qy 973 CTCTGGAGACCTTAATCTCATGC -- AGCCAAGTCGTCCTGGGTTAAAGAAAAA 1029
 Db 235 CTACTGACCCATCAAGCCCCTAGCGGAAAGAACCTGCGAACAGAAAG 176

Qy 1030 AAGGACAGTGGAAACATTGATCACCCTGATCSCGGTTCTACCCGAAAGGATGTCGGAGAC 1089
 Db 175 AAGGACAGTGGACACTGAAAGATGCTGTCGACCCGAAAGATGTCGACGCCAC 116

Qy 1090 ATGGTGTGACACUTGAAAGGACTCTCTGATTCCTCTGATTTGTTGACAGATTTGGG 56
 Db 115 CTGGTGTCCACACAGATGAAAGACTCTGTGCGAGTCTGCCCCAGAGTTGGG
 Db 55 CGCAAGGATCACCTCACCTCACCCGCAATACCAAGAACCCACTACAGAGCTGATGA 1204

RESULT 13
ABA22432/C
ID ABA22432 standard; DNA; 475 BP.

XX AC ABA22432;
 XX DT 23-JAN-2002 (first entry)
 DE Probe #898 for gene expression analysis in human heart cell sample.
 XX Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease; ss.
 OS Homo sapiens.
 XX PN WO200157274-A2.
 PN XX PD 09-AUG-2001.
 PN XX PF 30-JAN-2001; 2001WO-US00666.
 PN XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-020456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000US-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX PS Claim 1; SEQ ID No 898; 530pp; English.
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 475 BP; 75 A; 129 C; 142 G; 129 T; 0 other;

SQ Query Match Score 3.7%; Length 475;
Best Local Similarity 74.1%; Pred. No. 4.3e-41;
Matches 352; Conservative 0; Mismatches 120; Indels 3; Gaps 1;

Qy 733 ACTCATTCCTCGAGAAACCAACAAGTGTAAATTGAGAAATTTACCGAAA 792
Db 475 ACCATTTCTCCCGAAATCTCACCGTGTGCCTACTGTAGAACGGTCAA 416

Qy 793 GATCATCGAGAATCATCTCCATACAGAACCCCTAACAAACAGACCTTAAGTGCGAA 852
Db 415 GACCACTGAAACCTCCGAGCCACAAATGGCTTGGGTGTGA 356

Qy 853 GAATGTGGAGAAGACTCATACAGCTTGGATTAAACGTCACITGCCCTGCATCC 912
Db 355 GAGTGTGGAGAAGACTAACACCATGTGGCTTACAGGECACCTGCCATCG 296

Qy 913 GCARCAATGGTGGAGACTCCTGAAACTTGTGAAACCTTGTGAAAGCTGAGTC 972
Db 295 GCGAGTAGGGGAGCTTACCTGTTGGCTTACAGGECACCTGCCATCGT 236

Qy 973 CTTCCTGGAGCACCTTAATCTCATGC - - AGGAAGACTCTCTGTGGTTAAGAAAA 1029
Db 235 CTACTGGACCACTCAAGGCCATGAGGAGTAAACCCCTAACGGACCAAGAAG 176

Qy 1030 AAGCACCATGGGAGACATSTGTGATGCCGTTCTACACCGAAAGGATGTCGGAGAAC 1089
Db 175 AAGCACCATGGGAGACATSTGTGATGCCGTTCTACACCGAAAGGATGTCGGAGCAC 116

Qy 1090 ATGGGGTGTGACACTGGAAGAACCTTCTGTCAGATTTGGCACAGATTTGG 1149
Db 115 CTGGGGTCCACAGGATGCAAGGACTCTCTGTGCAAGTCTGTCGGAGATTGGG 56

Qy 1150 CGAAAGGATCACCTGACTCGACATATGAAAGAGTCACATCAAGGCCTCTGA 1204
Db 55 CGCAAGGATCACCTGACTCGACATATGAAAGAGTCACATCAAGGCATACAGAGGCTCTGA 1

RESULT 14
ID AAK00905/C
AC AAK00905;
XX 05-NOV-2001 (first entry)
DE Human brain expressed single exon probe SEQ ID NO: 896.
KW Human; brain expressed exon; gene expression analysis; probe;
microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
epilepsy; cancer; ss.
XX Homo sapiens.
XX WO200157275-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00667.
PF 04-FEB-2000; 2000US-01180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human brains -

XX Example 4; SEQ ID NO: 896; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.

XX Sequence 475 BP; 75 A; 129 C; 142 G; 129 T; 0 other;

SQ Query Match Score 3.7%; Length 475;
Best Local Similarity 74.1%; Pred. No. 4.3e-41;
Matches 352; Conservative 0; Mismatches 120; Indels 3; Gaps 1;

Qy 733 ACTCATTCCTCGAGAAACCCACAGCTGTAATTGAGAAATGTTACCGAAA 792
Db 475 ACCATTTCTCCCGAAATCTCACCGTGTCACTGTGAGAACGTGTCACCGAAA 416

Qy 793 ACTCATTCCTCGAGAAACCCACAGCTGTAATTGAGAAATGTTACCGAAA 792
Db 475 ACCATTTCTCCCGAAATCTCACCGTGTCACTGTGAGAACGTGTCACCGAAA 416

Qy 913 GATCATCGAGAAGACTCATACAGCTTGGATTAAACGTCACITGCCCTGCATCC 912
Db 235 GACCACTGAAACCTCCGAGCCACAAATGGCTTGGGTATAAGGGACCTTGTGCA 852

Qy 1030 AAGCACCATGGGAGACATSTGTGATGCCGTTCTACACCGAAAGGATGTCGGAGAAC 1089
Db 175 AAGCACCATGGGAGACATSTGTGATGCCGTTCTACACCGAAAGGATGTCGGAGCAC 116

Qy 1090 ATGGGGTGTGACACTGGAAGAACCTTCTGTCAGATTTGGCACAGATTTGG 1149
Db 115 CTGGGGTCCACAGGATGCAAGGACTCTCTGTGCAAGTCTGTCGGAGATTGGG 56

Qy 1150 CGAAAGGATCACCTGACTCGACATATGAAAGAGTCACATCAAGGCCTCTGA 1204
Db 55 CGCAAGGATCACCTGACTCGACATATGAAAGAGTCACATCAAGGCATACAGAGGCTCTGA 1

RESULT 15
ID AAK26350/C
AC AAK26360;
XX 06-NOV-2001 (first entry)
DE Human bone marrow expressed single exon probe SEQ ID NO: 917.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukaemia; lymphoma; ss.
XX Homo sapiens.
XX WO200157276-A2.
XX

PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001W0-US00668.
 XX
 PR 04-FEB-2000; 2000JS-0180312.
 PR 26-MAY-2000; 2000JS-0207456.
 PR 30-JUN-2000; 2000JS-0608408.
 PR 03-AUG-2000; 2000JS-0632356.
 PR 21-SEP-2000; 2000JS-0234687.
 PR 04-SEP-2000; 2000JS-0236359.
 PR 04-OCT-2000; 2000JSB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 917; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 475 BP; 75 A; 129 C; 142 G; 129 T; 0 other;

Query Match	Score	DB	Length
Best Local Similarity	3.7%	267	475;
Matches	74.1%;	Pred. No.	4.3e-41;
Db	0;	Mismatches	120;
QY	352;	Indels	3;
		Gaps	1;
Db	733	ACTCACTCTCTGAGAAACCCACAACACTGTAAATTATGTGAGAAATGTTTCACCGAAA	792
QY	475	ACCCATTCTCCCAGAAATCTACCGAGTGTCTACTGTGAGACGTAAACCGAAA	416
Db	793	GATCATCTGAGAAATCACCTCCATCACACGACCCCTAACAAAGAGACGTTTAAGTGCAA	852
QY	415	GACCACCTGAGAAACCCATTCCAGCCCCAACAAATGCCTTGGGTGTGAG	356
QY	853	GAATGTGGCAAGAACTACATAACCAAGCTGATTTAAACCTACCTGGCTTGCATGCC	912
Db	355	GAGTGTGGCAAGAACTACACCACTGCTGGCTATAAGGGCACTGGCTCATGCC	296
QY	913	GCAACAGTGGTGAACCTCACTGTAGGTATGTTGCAAACTTTGAAAGCACGGAGTG	972
Db	295	GCCAGCAGTGGGACCTCACCTGGGGCTGCTGGGTGTGGAGTAGGGACACCGAGTG	236
QY	973	CTTCCTGGAGCACCTTAATCTCATGC--AGGCAAGTCGCTGCTGGGGTTAAGAAA	1029
Db	235	CTACTGGACACCTCAAGGCCATGCGAGAGGCCCTACGGAACAAAG	176
QY	1030	AAGCACCAGTCGAACATCTGATGCGCTCTCACCCGAANGATTCGGAGAAC	1089
Db	175	AAGCACCAGTCGAACATCTGAAAGATGCTTCAACCGGAGGAATGCGACGCCAC	116
QY	1090	ATGGTGGTGAACACTGAAAGAAAGACTCTCTGTCAGATTGTCAGAAGATTTGG	1149
Db	115	CTGGTGGTCAACAGGATCAAGGACTCTCTGTCAGATTGTCAGAAGATTTGG	56
QY	1150	CGAAAGGATACCTGACTGACATGAAAGAAGACTACATGAAAGAGTCAGATCTGA	1204
Db	55	CGCAAGGATACCTGCCATACCAAGAACCCACTCACAGGAGTGTGATGATGA	1



FEATURES	Faraday Avenue Genoscope sequence ID : XCL0BB001ZD02RPL.	Qy	1493	CTACCTCATATGCCATTCTTCAATTCTGAAAGAACAGCCATTAAAGGGAAATTGAGA	1552
source	Location/Qualifiers 1. .1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="XCL0BB001ZD02" /issue_type="NEUROBLASTOMA" /clone.Lib="Homo sapiens NEUROBLASTOMA" /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dtr) primer. Five prime end enriched double-strand cDNA was digested with Not I and cloned into the Not I and EcoR sites of the pCMVSPORT 6 vector. Library was not normalized." BASE COUNT	Db	1088	CTACCTCATATGCCATTCTTCAATTCTGAAAGAACAGCCATTAAAGGGAAATT-	1145
	RESULTS	Qy	1553	GTTACTGTGAGGTACAGGTGCTCCTCATCCAAAGATTCTCAAGGATC	1610
	LOCUS	Db	1146	RKTACCGGATGGACTACAAGKGGC-TKCCCTCTCATCCC-ARATWTCAGCAT	1201
	DEFINITION				
	Homo sapiens cDNA clone CS001012M24 5'-PRIME, mRNA sequence.				
	ACCESSION	BX370812			
	VERSION	BX370812.1			
	KEYWORDS	GI:30445558			
	SOURCE				
	ORGANISM				
	Homo sapiens (human)				
	ORGANISM				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.				
	REFERENCE				
	1 (bases 1 to 866)				
	AUTHORS				
	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
	TITLE				
	Full-length cDNA libraries and normalization				
	JOURNAL				
	Unpublished				
	COMMENT				
	Contact: Genoscope - Centre National de Sequencage				
	Genoscope - France				
	BP 191- 91006 Evry cedex - France				
	Email: seqref@genoscope.cnrs.fr. Web: www.genoscope.cnrs.fr				
	Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4473.r. For more information about this cluster, see				
	cgi-bin/cluster.cgi?seq=CS0BAG062ZD02_CS05868_1&cluster=4473.r.				
	Contact : Feng Liang Email : fliang@tifetch.com URL :				
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BAG062ZD02_CS05868_1.				
	FEATURES				
	SOURCE				
	organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS0D1012M24"				
	/cell_type="B CELLS (RAMOS CELL LINE)"				
	/cell_line="RAMOS CELL LINE"				
	/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT				
	25-NORMALIZED"				
	/note="1st strand cDNA was primed with a NotI-oligo (dtr) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the PCMVSPORT 6 vector. Library was normalized."				
	BASE COUNT	235 a	218 b	177 g	236 t
	ORIGIN				
	Query Match				
	Best Local Similarity				
	Matches				
	843;	Conservative			
	0;	Mismatches			
	3.1e-72;	Indels			
	4;	Gaps			
	Qy	984	CCCTAAATCTCATGCCATTCTGAGGAAAGGATCTGGTGGGAAAGCACAGTGGGA	1043	
	Db	14	CCCTAAATCTCATCA-GGAAGCTGGTGGGTTAAAGAAAAGGC-CCAGTGGGA	71	
	Qy	1044	ACATTGTGATCGGGCTTACCCGAAAGGATCTGGTGGGAAAGCACATGGTGTGCAAC	11.03	
	Db	72	ACATTGTGATCGGGCTTACCCG-AAGAGTGGTGGGAGACATGGTGTGCAAC	130	
	Qy	1104	TGGAAAGGACTATGAAAGGAGTCAATAAGGACTCTGAGGTTCTGAGGTCACCT	11.63	
	Db	131	TGGAAAGGACTCTGGTGTGCAATGGTGGGAAAGGATCACCT	190	
	Qy	1164	GACTGCACATATGAAAGGAGTCAATAAGGACTCTGAGGTTCTGAGGTCACCT	12.23	
	Db	191	GACTGCACATATGAAAGGAGTCAATAAGGACTCTGAGGTTCTGAGGTCACCT	250	

FEATURES	source	Query Match Best Local Similarity 95.4%; Matches 759; Conservation 0;	Score 688; DB 9; Length 790; Pred. No. 2e-61; Indels 9; Gaps 5;	Db	773 ACRAAACTCAA 784
FEATURES	Location/Qualifiers	1..790	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="PLACE1008574" /tissue_type="placenta" /clone_1is="PLACE1" /note="Vector: PME18SFL3"		
BASE COUNT		241 a 138 c 140 g 267 t	4 others		
ORIGIN					
Query Match Best Local Similarity 95.8%; Matches 759; Conservation 0;	Db	2190 GTCCCTGATTCAGTGCCTAAATTGCAATTGAGACTTTTATAGGGAAAGC 2249	Score 688; DB 9; Length 790; Pred. No. 2e-61; Indels 9; Gaps 5;	Db	CA413047/c
	Qy	1 GTCCGTGATTCAGTGCCTAAATTGAGACTTTTATAGGGAAAGC 57		DEFINITION	UI-H-EZ0-bap-e-03-0-UI-s1 NCI CGAP Ch1 Homo sapiens mRNA sequence.
	Db	2250 CATTATTAGTAGTAACTTACAAATCCATTCAAAATTACCTTTAGATCTAAATT 2309		ACCESSION	CA413047
	Qy	58 CATTATTAGTAGTAACTTACAAATCCATTCAAAATTACCTTTAGATCTAAATT 117		VERSION	CA413047.1
	Db	2310 TTCAATTTCCTCTATAAACAGTGGCTCTAACCTTTGACACTCTGCTCATTA 2369		KERTON	EST.
	Qy	118 TTCAATTTCCTCTATAAACAGTGGCTCTAACCTTTGACACTCTGCTCATTA 177		SOURCE	Homo sapiens (human)
	Db	2370 GCAATAGAATGTAATTGTATAAAAGTTGTGATAACTCAAGGTAAATTCTTA 2429		ORGANISM	Homo sapiens
	Db	178 GCAATAGAATGTAATTGTATAAAAGTTGTGATAACTCAAGGTAAATTCTTA 237		COMMENT	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominidae; Homo.
	Qy	2430 CTAGCTTCTTAATGGTTATAATCAAGGCTTCAAATTAAGCTTCAAGTTTCGGA 2489		REFERENCE	1 (bases 1 to 742)
	Db	238 CTAGCTTCTTAATGGTTATAATCAAGGCTTCAAATTAAGCTTCAAGTTTCGGA 297		AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
	Qy	2490 AGATAATAATGTTGTTGATAACCCATAATTTCAGATCAGTATTCGAGACTTC 2549		TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
	Db	298 AGATATAATGTTGTTGATAACCCATAATTTCAGATCAGTATTCGAGACTTC 357		JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgaps-x@mail.nih.gov
	Qy	2550 TGTGTGCTGGCTAAATAATTTCGATCTTATTATGAGCTTAAAGAAAACCTTA 2609		ORTHOGENICS	Orthopaedics
	Db	358 TGTGTGCTGGCTAAATAATTTCGATCTTATTATGAGCTTAAAGAAAACCTTA 417		CDNA LIBRARY PREPARATION	CDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
	Qy	2610 AACACAAGCATCAGTATTATAGCAAAGAGACTCTGTAGGTGACATGGCATTCG 2669		TISSUE PROCUREMENT	Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of Orthopaedics
	Db	418 AACACAAGCATCAGTATTATAGCAAAGAGACTCTGTAGGTGACATGGCATTCG 477		DNA SEQUENCING	CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
	Qy	2670 GTCACTTAATAGTGGCTTAATTAGTACACAGGATATTGTCTGTTCATCTTC 2729		CLONE DISTRIBUTION	Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
	Db	478 GTCACTTAATAGTGGCTTAATTAGTACACAGGATATTGTCTGTTCATCTTC 537		THE FOLLOWING REPETITIVE ELEMENTS WERE FOUND IN THIS cDNA	The following repetitive elements were found in this cDNA sequence: 1-21, >AT rich<Low complexity (matched compliment
	Qy	2730 TAACATGCTTAATCTTCATTTAATAT-AGTAATGTTGTTGCAATGCTTC 2788		SEQ PRIMER	Seq primer: M13 FORWARD
	Db	538 TAACATGCTTAATCTTCATTTAATAT-AGTAATGTTGTTGCAATGCTTC 597		POLY-A/YES	POLY-A/YES.
	Qy	2789 TCGATATATAGTGGTATCTTCTTCACTTAAATGTTGCAATGCTTCAGAG 2848		FEATURES	Location/Qualifiers
	Db	598 TGATGATATAGTGGTATCTTCTTCACTTAAATGTTGCAATGCTTCAGAG 657		SOURCE	1. 742
	Qy	2849 GTGCGAGCCAGAAATTATTTCATTTCTCATGGTGGCCACGGTTAGTCATTAGGCCATT 2908		ORGANISM	/organism="Homo sapiens"
	Db	658 GTGCGAGCCAGAAATTATTTCATGGTGGCCACGGTTAGTCATTAGGCCATT 716		MOL TYPE	/mol_type="mRNA"
	Qy	2909 GTGGAACTGGAGCCATGAGGTGGTTAACCTTCAAGTGGTTGACTCCTCTGTGGT 2968		DB XREF	/db_xref="ITaxon:9606"
	Db	717 GTGGAACTGGAGCATGAGGTGGTTAACCTTCAAGTGGTTGACTCCTCTGTGGT 772		TISSUE	/tissue="Chondrosarcoma Grade II"
	Qy	2969 GACAGACTTCTA 2980		DEV STAGE	/dev_stage="Adult"
	Qy	6596 GTATTCTATGTGACTTTGGAAAGGAGTCTCGGACATGCAANGTTACGACAAACC 6655		LAB HOST	/lab_host="DH10B (Life Technologies)"
	Db	742 GTATTCTATGTGACTTTGGAAAGGAGTCTCGGACATGCAANGTTACGACAAACC 684		CLONE LIB	/clone_lib="NCI CGAP Ch1"
	Qy	6656 AGCTTTGTGCCCCAC-AAGGTGACTGTAACTCAGATGAAAGTGGCTTATAATGGCTG 6714		NOTE	/note="Organ: Left Pelvis; Vector: PT773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch1 is a cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into PT773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag (dT)18 tail. The sequence tag for this library is TAG-LIB-UU-H-EZ0
				TAG TISSUE	TAG-PRATE-2-chondrosarcoma
				BASE COUNT	236 a 102 g 274 t 1 others
				ORIGIN	

us-09-242-772-116.rst

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). Seq Primer: M13 Reverse.	
FEATURES	source
683 AGCTTGTCCNAAAGGTGACTGTAATCAGAATGGCTTTATAATAGGGTG 624	/organism="Homo sapiens" /mol_type="mRNA"
6715 TGGAGTGAAGAACATGCTGATGTTACTAACAGCCPTTGAAATTACAAAACTGGAA 6774	/db_xref="MGI:9606"
623 TGGAGTGAAGAACATGCTGATGTTACTAACAGCCPTTGAAATTACAAAACTGGAA 564	/clone="UI-E-EO1-ii-z-j-24-0-UI"
6775 TCATTAGAAAGGGATGCACTACCTGAACTGCTGACTGCTGAAATTGTATT 6834	/tissue_type="fetal eye"
563 TCATTAGAAACCGATTGCACTACCTGAACTGCTGAAATTGTATT 504	/dev_stage="fetal"
6835 TTAGCTAATGAAAGGTGGACTGACTACTCTAAATGTTCTAATGATAAAAGTTTG 6894	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
503 TTAGCTAATGAAAGGTGGACTGACTCTAAATGTTCTAATGATAAAAGTTTG 444	/clone_lib="UI-E-EO1"
6895 AGTCAAAATGAAAGRAAAAATGCAATTGCAATTGTTGATTTTATGCC 6954	/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EO1 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded DNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dR)18 tail. The sequence tag for this library is CGCTATACC. This library was created for the Program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
443 AGTCAAAATGAAAGRAAAAATGCAATTGCAATTGTTGATTTTATGCC 384	BASE COUNT
6955 ATTTAAAATGCTATTCTGTAATATTGGAAATCAAGTCACTGTTATCGTGA 7014	ORIGIN
323 CTTAAAATGATTCAAACACTGTGTTATGATAATAGACAAGATATAATT 264	Query Match 9.4%; Score 685; DB 12; Length 693;
7075 TTGTTGTTGTGTATAAGCTGTGAAATAAGGAAATCATATGAGCTGAGTGACTA 7134	Best Local Similarity 100.0%; Pred. No. 4.4e-61;
263 TTGTTGTTGTGTATAAGCTGTGAAATAAGGAAATCATATGAGCTGAGTGACTA 204	Matches 0; Mismatches 0; Indels 0; Gaps 0;
7135 CATTITAAACATTCACATCCAAGGCAAGCTATTTCGTCATATACCAAGATTTAA 7194	51133 TGAGTGTGGAAAGGATCTTGGACTGTAGCANTGCTTCCATTGATGTTATAAGTACCT 5192
203 CATTITAAACATTCACATCCAAGGCAAGCTATTTCGTCATATACCAAGATTTAA 144	Matches 685; Conservative 0;
7195 ATATTAAATTGTGTGTATAATTAAGTACTGCGAGCTTCTTGTGCCTCACAGTTT 7254	Qy 9 TGAGTGTGGAAAGGATCTTGGACTGTAGCANTGCTTCCATTGATGTTATAAGTACCT 68
143 ATATTAAATTGTGTGTATAATTAAGTACTGCGAGCTTCTTGTGCCTCACAGTTT 84	Db 5193 TTATAGTTATCGATCACTGTGTTAAACACTTCATTTTAAATCTTATACCAAGTTGAGTT 5252
7255 GTTTGCTGTAAAGAATGATGATAATGTGAATTCCAAATAATGAAATAAT 7305	Qy 69 TTATAGTTATCGATCACTGTGTTAAACACTTCATTTTAAATCTTATACCAAGTTGAGTT 128
83 GTTTGCTGTAAAGAATGATGATAATGTGAATTCCAAATAATGAAATAAT 33	Db 5253 TTAAAACTTCATTGCTCTGGTATGCACTCTGCTGCAACTTTTATTTC 5312
7 6M724590 693 bp mRNA linear EST 01-MAR-2002	Qy 129 TTAAAACTTCATTGCTCTGGTATGCACTCTGCTGCAACTTTTATTTC 188
DEFINITION UI-E-EO1-aii-z-j-24-0-UI.r1 UI-E-EO1 Homo sapiens cDNA clone	Db 5313 ATTAGTGTGTTCTTCAAGCTGTATTTGCCTATTGCTGCTTATTTC 5372
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata, Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.	Qy 189 ATTAGTGTGTTCTTCAAGCTGTATTTGCCTATTGCTGCTTATTTC 248
ACCESSION BM724590-1	Db 5373 TTAGTCATTGTTGGAAATATAGTGTATAATTGTGTAAATTGGCAATTTTTAA 5432
VERSION GI:119045921	Qy 249 TTAGTCATTGTTGGAAATATAGTGTATAATTGTGTAAATTGGCAATTTTTAA 308
KEYWORDS Homo sapiens (human)	Db 5433 AACATTATACTGACTGAAACATGACCCGAAAGCCGATGCTTATTGCTGAACTTTTATTTC 5492
ORGANISM Homo sapiens	Qy 309 AACCTTATACTGACTGAAACATGACCCGAAAGCCGATGCTTATTGCTGAACTTTTATTTC 368
DEFINITION 1 (bases 1 to 693)	Db 5493 GTTAAGAGTACATTTCAGATGTTCTATGAGCAATCATATTAGGCC 5555
REFERENCE 1 (bases 1 to 693)	Db 369 GTTAAGAGTACATTTCAGATGTTCTATGAGCAATCATATTAGGCC 428
AUTHORS Bonaldo, M.F.; Lennon, G. and Soares, M.B.	Qy 5553 AATTTTTTTAAATAGAGCTTGTCAACCTCTATACTACATAATTACAGATATA 5612
TITLE Normalization and subtraction: two approaches to facilitate gene discovery	Db 5554
JOURNAL Genome Res. 6 (9), 791-806 (1996)	Qy 5613 GCACTTCAAAATGAACTTACAGCTTCTATACATACATAATTACAGATATA 5677
MEDLINE 97044477	Db 429 AATTTTTTTAAATAGAGCTTGTCAACCTCTATACTACATAATTACAGATATA 488
PUBMED 8889548	Qy 489 GCACTTCAAAATGAACTTACAGCTTCTATACATACATAATTACAGATATA 548
COMMENT Coordinated Laboratory for Computational Genomics	
University of Iowa	
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA	
Pax: 319 335 8250	
Email: bentosoares@uiowa.edu	
Tissue Procurement: Dr. Gregg Hageman	
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa	

Qy	5673	TTAGACTTATAATTCAAGTCCATTAGATGATAATATGAGCTTGATATA	5732	Qy	181	GGCGTATTGGACCCGAATGCCCTATTCTGACGGTCTATTAAAGCTGCCCTG	240
Db	549	TTAGACTTATAATTCAAGTCCATTAGATGATAATATGAGCTTGATATA	608	Db	274	GGCGTATTGGACCCGAATGCCCTATTCTGACGGTCTATTAAAGCTGCCCTG	333
Qy	5713	ATGCTTGTGTGTCAAAATGGTATGGTATTTAACAGGTACATTTACAGTGT	5792	Qy	241	CTAGAAGTTGGACAGGGCCCTAGATGGCAAATATGGAAAGATGGATTCACCTCT	300
Db	619	ATGCTTGTGTGTCAAAATGGTATGGTATTTAACAGGTACATTTACAGTGT	668	Db	334	CTAGAAGTTGGACAGGGCCCTAGATGGCAAATGGAAAGATGGATTCACCTCT	393
Qy	5793	CCTATCATTGCTATATGCACAG	5817	Qy	301	TCCAGAAGAGTCATGGACTGGCTAGATCAAAGTCGAGCTTTCCATAGTAAT	360
Db	669	CCTATCATTGCTATATGCACAG	693	Db	394	TCCAGAAGAGTCATGGACTGGCTAGATCAAAGTCGAGCTTTCCATAGTAAT	453
RESULT 8							
LOCUS	BX402039	1201 bp	mRNA	linear	EST 13-MAY-2003		
DEFINITION	BX402039	Homo sapiens B CELLS (RAMOS CELL LINE)	COT 25-NORMALIZED				
ACCESSION	BX402039	Homo sapiens cDNA clone CS0DL012YM24	5'-PRIME, mRNA sequence.				
VERSION	BX402039.1	GI: 30607215					
KEYWORDS	EST.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.							
REFERENCE	1 (bases: 1 to 1201)						
AUTHORS	Li, W. B., Gruber, C., Jessee, J. and Polayes, D.						
TITLE	FULL-length cDNA libraries and normalization						
JOURNAL	Unpublished						
COMMENT	Contact : Genoscope						
Genoscope - Centre National de Séquençage	BP 191 91006 EVRY Cedex - France.						
Email: seqref@genoscope.cns.fr. Web : www.genoscope.cns.fr							
was normalized. Library was constructed by Life Technologies a division of Invitrogen. This sequence belongs to sequence cluster 4473 . For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DL012BG12QPI1 .							
PENG LIANG Email : fliang@lifeitech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DL012BG12QPI1.							
FEATURES	Location/Qualifiers						
source	1..1201						
/organism="Homo sapiens"							
/mol_type="mRNA"							
/db_xref="Taxon:9606"							
/clone="CS0DL012YM24"							
/cell_type="B CELLS (RAMOS CELL LINE)" COT 25-NORMALIZED"							
/cell_line="RAMOS CELL LINE"							
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"							
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double stranded cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."							
BASE COUNT	319 a	260 c	297 g	273 t	52 others		
ORIGIN							
Query Match	9.1%	Score 666; DB 13; Length 1201;					
Best Local Similarity	80.4%	Pred. No. 2.8e-59;					
Matches	918; Conservative 6;	Mismatches 6; Indels 212; Gaps 4;					
Qy	1	GGCAGGGCATACACTACAAATGGCTGTGAAAGAGGGTAGGAAACAAATTCCAGGCC	60	Qy	1020	TAAGAAAAAAAGCACAGCTGCGAACATTGTGCGTGTGAAACCTTGAA	1079
Db	94	GGCAGGGCATACACTACAAATGGCTGTGAAAGAGGGTAGGAAACAAATTCCAGGCC	153	Db	905	TAAGAAAAAA-CACCACTGCGAACATTGTGCGTGTGAAACCTTGAA	963
Qy	61	GCGCSTCCAGCCGAATATGAGAAAAATTATGAGAAATTCCGGCCGGCTAGAG	120	Qy	1080	CCGGAGACACATGGTGTGAACTCTCTGTGAAAGCAAGTCGCTGCA	1139
Db	154	GCGCCTCCAGCCGAATATGAGAAAAATTATGAGAAATTCCGGCCGGCTAGAG	213	Db	964	CCGGAGACACATGGTGTGAACTCTCTGTGAAAGCAAGTCGCTGCA	1022
Qy	121	GCGGGGACCGGCCGAGGGAGGATGTTAAAGGCCCGCGGTTGCCTCTTGTGTG	180	Qy	1140	GA 1141	
Db	214	GCGGGGACCGGCCGAGGGAGGATGTTAAAGGCCCGCGGTTGCCTCTTGTGTG	273	Db	1023	GA 1024	
				RESULT 9			
				AL036879			
				LOCUS			

DEFINITION	DKFZp564P1863_r1_564 (synonym: htbr2) Homo sapiens cDNA clone DKFZp564P1863_5', mRNA sequence.	Qy	1179 GAAGAGTCACATCAAGAGTCATAAGGTCAAAAGAACCGAGTGGATTTCCTGCC 123.8
DB	422 GAGAGTCACATCAAGAGTCATAAGGTCAAAAGAACCGAGTGGATTTCCTGCC 48.1	Db	422 GAGAGTCACATCAAGAGTCATAAGGTCATAAGGTCATAAGAACCGAGTGGATTTCCTGCC 48.1
ACCESSION	AL036879	Qy	1239 ATTACCTGCAATGTCATGCTGCGCTATAAAGAGCAGGTCCCTCCGGTCAAATGCAATGCTTGCCT 129.8
VERSION	1	Db	482 ATTACCTGCAATGTCATGCTGCGCTATAAAGAGCAGGTCCCTCCGGTCAAATGCAATGCTTGCCT 54.1
KEYWORDS		Qy	1299 TTCCAGTG-AACTGTTATCAAACCATTCAACAAACCTCTAACACA 135.7
SOURCE	Homo sapiens (human)	Db	542 TTCCAGTGAAACTGTTATCAAACCATTCAACAAACCTCTAACACA 60.1
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 736)	Qy	1358 CTCCATTTCAGTCATGC-AGAGTCGGATCGCCACC-AGACACTGTCATCCCCTCACACCCTTCCT 141.5
REFERENCE	Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann S.	Db	602 CTCCATTTCAGTCATGCAGGTGGATGCCAACAAATGATCACACTTAC 66.1
AUTHORS	EST (Duesterhoeft, et al.)	Qy	1416 TTGGGAATGACATGCCAATAGATAAGTAT-GGACACTGTCATCCCCTCACACCCTTCCT 147.4
JOURNAL	Unpublished	Db	662 TTGGGAATGACATGCCAATAGATAAGTATGGACACTGTCATCCCCTCACACCCTTCCT 72.2
COMMENT	On Jul 7, 1999 this sequence version replaced gi:5866322.	Qy	1475 TCAAATATCGTTAGT 149.1
Contact:	Duesterhoeft A	Db	722 - -CATAATCGTTAGT 73.6
MIPS	Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany	RESULT	097 bp tRNA linear EST 18-APR-2001
	This is the 5' sequence of the clone insert	LOCUS	BG613583 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4772744 5'
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;	DEFINITION	mRNA sequence.
	Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;	ACCESSION	602641935FL NIH_MGC_61
	consorciated by Qigen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.	VERSION	BG613583
	No 5' sequence available.	KEYWORDS	GI:13664954
	This clone (DKFZp564P1863) is available at the RZPD in Berlin.	SOURCE	Homo sapiens (human)
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	ORGANISM	Homo sapiens
	Location/Qualifiers	COMMENT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	1..736	REFERENCE	1 (bases 1 to 907)
	/organism="Homo sapiens"	AUTHORS	NIH-MGC http://mgc.nci.nih.gov/
	/mol type="mRNA"	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
	/db_xref="taxon:9606"	JOURNAL	Unpublished
	/clone="DKFZp564P1863"	COMMENT	Contact: Robert Strausberg, Ph.D.
	/tissue type="brain"	EMAIL	cgaps-x@mail.nih.gov
	/dev_stage="fetal"	Tissue	Procurement: ARCC
	/lab_host="X1-blue"	Procurement	CLONETECH Laboratories, Inc.
	/clone lib="564 (synonym: hfb52)"	CDNA Library Preparation	Inc.
	/note="Vector: PAMPL; Site 1: NotI; Site 2: SalI"	CDNA Library Arrayed By	The I.M.A.G.E. Consortium (ILNL)
		CDNA Sequencing By	DNA Sequencing by Incyte Genomics, Inc.
		Clone Distribution	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
		http://image.llnl.gov	Plate: LICM1643 row: f column: 09
			High quality sequence S50P: 645.
		FEATURES	Location/Qualifiers
		Source	1..907
			/organism="Homo sapiens"
			/mol type="mRNA"
			/db_xref="taxon:9606"
			/clone="IMAGE:4772744"
			/tissue type="embryonal carcinoma"
			/lab_host="DH10B (T1 phage-resistant)"
			/clone lib="NTH_MGC_61"
			/note="Organ: testis; Vector: pDNRL-LIB (Clontech); Site_1: SF1 (ggccgcgtccggccc); Site_2: SII (ggcccatatggcc); Site_3: adaptor sequence: 5'-ATTCCTAGGGCCAGGGCCAGCATG-dT(30)BN-3', where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT	267 a	202 c	174 g	264 t
ORIGIN				
Query Match	8.6%	Score 629; DB 10;	Length 907;	
Best Local Similarity	87.7%	Pred. No. 1 ge-55;		
Matches 774; Conservative	0; Mismatches 85; Indels 24; Gaps 7;			
Db	2763 AGTGTATGGCATTTGGGCTTCAGTGATAATAGGTAGATCTCTTAGTCCTTCA 2822			
Db	1 ACTGTATGGCATTTGGGCTTCAGTGATAATAGGTAGATCTCTTAGTCCTTCA 60			
Qy	2823 CCTTTACATCCAAAGGCGGGTCAACTCAGCAGGATTATTTCTATGTGGCCA 2882			
Db	61 CCTTTACATCCAAAGGCGGGTCAACTCAGCAGGATTATTTCTATGTGGCCA 120			
Qy	2883 CGGTAGTCATAATCTAGACCATTGTGACTGAGCCTAGCTGAGGTGTATCCAC 2942			
Db	121 CGGTACTCATATACTAGACCATTGTGACTGAGGTGTATCCAC 180			
Qy	2943 AGTGTATTGACTCAGCTCTGTGGTCAAGCAGCTTAAGCGGAAAGATAAGCTGAAGC 3002			
Db	181 AGTGGATTGACTCAGCTCTGTGGTCAAGCAGCTTAAGCGGAAAGATAAGCTGAAGC 240			
Qy	3003 ACATGGTTAACATTGGGAACTTGTGTAGGGATCATGCCCTGTAGCCGGTTAAAC 3062			
Db	241 ACATGGTTAACATTGGGAACTTGTGTAGGGATCATGCCCTGTAGCCGGTTAAAC 299			
Qy	3063 TGGACTTTTATAGAAGTAAAGCTAAAGCATAGGCCTTATATCATTCTGTGAATTGTAT 3122			
Db	300 TGGACTTTTATAGAAGTAAAGCTAAAGCATAGGCCTTATATCATTCTGTGAATTGTAT 359			
Qy	3123 ATGGTTTTCTTCCCTTAAGAATCAAACAAACAAAGGAGAAACAAACAGTCCTCATCCGAT 3182			
Db	360 ATGGTTTTCTTCCCTTAAGAATCAAACAAACAAAGGAGAAACAAACAGTCCTCATCCGAT 419			
Qy	3183 GTPATCTTGTGATTCAATGGAATCCATTTCTGAAATATTGTTGATGGACAAATT 3242			
Db	420 GTPATCTTGTGATTCAATGGAATCCATTTCTGAAATATTGTTGATGGACAAATT 479			
Qy	3243 GAACCTTAATGCAATTGAAAACCTACCTGATGAACCTACCTGATGAACTTAACTGAGA 3302			
Db	480 GAAGTTAAATGCAATTGAAAACCTACCTGATGAACTTAACTGAGA 539			
Qy	3303 ATATATACAGTAAATCCCACTTCTGATCTGGAAATGATGAGGATTGCAAT 3362			
Db	540 ATATATACAGTAAATCCCACTTCTGATCTGGAAATGATGAGGATTGCAAT 599			
Qy	3363 AAGTTGAGTTGTAGGGTAACAAAGTAAGTAAACCTATCTGGTAAACAT - 3420			
Db	600 AACCTGGAG-TGTAGGGTAACAAAGTAACCTATCTGGTAAACAT - 658			
Qy	3421 -GAAAATAACATTGAGATAATT-ATATTCATGTAAATAATT---AGCTTCTCT 3474			
Db	659 ACCATTCACATTGAGATAATTCTCATTCATGTAACTCTCTGGCTTTCCT 718			
Qy	3475 CACATTAGACACCAACATACTCTTAAAGT-----CTAATTAATATTT 3523			
Db	719 CCACATTAGACACCAACACTAGTCCTCTTACAGGTCTACTCAATCTCATCC 778			
Qy	3524 TCTCTAAGGGTCAAGTGGGACATAAC---CTAAGAACATCTATAAGCCTTAA 3580			
Db	779 TCTCTAAGGGTCAAGTGGGACATAACCTCTACATACCTAAGGCTCTGTAA 838			
Qy	3581 CACCTTATTAGGACACCAACATACTCTTCAAGGATGGGGCAAGGG 3623			
Db	839 ACCTTATTAGGACACCAACACTAGTCCTCTTACAGGTCTACTCAATCTCATCC 881			
Qy	6637 ATGTTACAGAAACAACCCAGTTTGTGTCACAGGTGACTGAACTCAGAATGGAAAGT 6696			
Db	661 AGCTTACAGAAACAACCCAGTTTGTGTCACAGGTGACTGAACTCAGAATGGAAAGT 552			
Qy	6697 GGCTTACAGAAACAACCCAGTTTGTGTCACAGGTGACTGAACTCAGAATGGAAAGT 6756			
Db	551 GGCTTACAGAAACAACCCAGTTTGTGTCACAGGTGACTGAACTCAGAATGGAAAGT 492			
RESULT 11				
BU686683/c	BU686683 1	EST	GI:23541826	
LOCUS	616 bp	mRNA	linear	EST 07-Oct-2002
DEFINITION	UI-CF-DU1-adp-o-08-0-UI.s1	UI-CF-DU1	Homo sapiens cDNA clone	UI-CF-DU1-adp-o-08-0-UI 3', mRNA sequence

Db	491	TTCACAAACTGGATCCATTAGAACCGATTGATCATACCTGAACTAAGTGG	432		made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction.
Qy	68117	ACTGCTGAATGTTATTGTGTAATGAAAGTGTGGACTTAGTACTCTAAATGT	6876	The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clone IDs 1414920-1417991 and 1520304-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."	
Db	431	ACTGCTGAATGTTATTGTGTAATGAAAGTGTGGACTTAGTCTAAATGT	372		
Qy	68177	TCAATGATAAAGTTTGAGTCAAATGAAAGAAAATCTGCATTCAGGCCGAT	6936		
Db	371	TCTAATGATAAAGTTTGAGTCAAATGAAAGAAAATCTGCATTCAGGCCGAT	312	BASE COUNT	
Qy	6937	TTCGTATATTTTATTGCAATTAAATGAAATGGAAATCAGTGGCT	6996	ORIGIN	
Db	311	TTCGTATATTTTATTGCAATTAAATGAAATGGAAATCAGTGGCT	252	Query Match	
Qy	6997	TATCATGTTATATCGTGTACTTAAATGATTCAACTACTGTTGTTGATAAAAT	7056	Best Local Similarity	
Db	251	TATCATTATTATCGTGTACTTAAATGATTCTGTTATGGAAATGTTGATAAAAT	192	0.100.0%	
Qy	7057	ATAGCAACATCATATTTTGTGTGTATANGCTGTAAATAGCAATCACATAT	7116	Pred. No. 2.7e-49;	
Db	191	ATAGACAAAGATCATATTGTTGTGTGTATAAGCTCTAAATAGCAATCACATAT	132	Mismatches	
Qy	7117	GAAGCTGCACTGATAACATTCACATCCAAAGAACGACTTATTTGT	7176	0	
Db	131	GAAGCTGCACTGATAACATTCACATCCAAAGAACGAGCTTATTTGT	72	Indels	
Qy	7177	CCATATACCAGATTAAATATTAAATTGTGTCTAAATATAATAGTACTGCA	7230	Gaps	
Db	71	CCATATACCAGATTAAATATTGTGTCTAAATATAATAGTACTGCA	18	0	
RESULT 12					
LOCUS	AW339499	570 bp mRNA linear EST:31-JAN-2000			
DEFINITION	xz91905.x1 NCI_CGAP_Lu24	Homo sapiens cDNA clone IMAGE:2871608 3',			
mRNA sequence.					
ACCESSION	AW339499				
VERSION	AW339499.1	GI:683125			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.			
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
TITLE	Unpublished				
JOURNAL	Contact: Robert Strausberg, Ph.D.				
COMMENT	Email: cgs@nlm.nih.gov				
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.				
	cDNA Library Preparation: M. Bento Soares, Ph.D.				
	CDNA Library Arrayed by: Greg Lennon, Ph.D.				
	DNA Sequencing by: Washington University Genome Sequencing Center				
	Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLM at:				
	www-bio.llnl.gov/bibr/image/image.html				
	Seq primer: -40UP from Gibco				
	High quality sequence stop: 471.				
FEATURES	Location/Qualifiers				
source	1..570	/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/db_xref="taxon:3606"			
		/clone IMAGE:2871608"			
		/tissue_type="carcinoind"			
		/lab_host="DRI0B"			
		/clone_id="NCI_CGAP_Lu24"			
		/note="Organ: Lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were			
RESULT 13					
LOCUS	BG253502	922 bp mRNA Homo sapiens cDNA clone IMAGE:4472455 5',			
DEFINITION					
mRNA sequence.					
ACCESSION	60236227F1	NIH_MGC_90			
VERSION	BG253502	linear EST:13-FEB-2001			
KEYWORDS	1 (bases 1 to 922)				
EST.	NIH-MGC	http://mgc.nci.nih.gov/			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished				
COMMENT	Contact: Robert Strausberg, Ph.D.				
	Email: cgs@nlm.nih.gov				
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.				
	cDNA Library Preparation: M. Bento Soares, Ph.D.				
	CDNA Library Arrayed by: Greg Lennon, Ph.D.				
	DNA Sequencing by: Washington University Genome Sequencing Center				
	Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLM at:				
	www-bio.llnl.gov/bibr/image/image.html				
	Seq primer: -40UP from Gibco				
	High quality sequence stop: 471.				
FEATURES	Location/Qualifiers				
source	1..570	/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/db_xref="taxon:3606"			
		/clone IMAGE:2871608"			
		/tissue_type="carcinoind"			
		/lab_host="DRI0B"			
		/clone_id="NCI_CGAP_Lu24"			
		/note="Organ: Lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were			

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
cDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC
Clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at:

Qy	6270	GAAGGCAGTGTCTTT	6284
Db	655	GAAGGCAGGGCCTT	669

RESULT	14
LOCUS	AGI12578/C
DEFINITION	pan troglodytes DNA, clone: RP43-016A13.TU, genomic survey sequence.
ACCESSION	AGI12578
VERSION	AGI12578.1
KEYWORDS	GSS .
SOURCE	Pan troglodytes (chimpanzee)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE	1
AUTHORS	Fujiiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE	BAC end sequences of Library RPCI-43
JOURNAL	Unpublished

AUTHORS	Fujiyama A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
TITLE	Direct Submission
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Saitoh-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/)
COMMENT	Clones are derived from the chimpanzee BAC library RPCI-43. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS	
SEQUENCING:	TU
LIBRARY	
Vector	: PBACE3.6
R.Site 1	: ECORI
R.Site 2	: ECORI
Location/Qualifiers	
source	1..659
/organism="Pan troglodytes"	
/mol_type="genomic DNA"	
/db_xref="taxon:9598"	
/clone="RP43-016A13.TU"	
/sex="male"	
/cell_type="lymphocytes"	
/clone_id="RPCI-43 Chimpanzee Male BAC Library"	
BASE COUNT	207 a 122 c 96 g 232 t 2 others
ORIGIN	
Query Match	7.6%
Best Local Similarity	9.45%
Matches	598; Conservative 0; Mismatches 32; Indels 3; Gaps 2;
Y	
6652	ACCCAGTTTGTCCACAAAGGTGACTCTGAATGGAAAGTCGGCTTTATAATAGG
6556	AACCAGTTTGTCCCGGGCTGACTCAAAATGGAAACTGCCTTATAATAGG
67112	GTCGGACTGAAACATGCTGTATGTACTACAGCCTTGAATTACAAAAAC-TG
596	GTCGGACTTAAAGAACATGCTGTATGTAGTACAGGCCCTTGAAATTACAAAAAC-TG
6771	GATCCCTTAGGAAACGGATGCAATACCTGAACATAAGCTGACTGCTGAAATTCT
536	GACATCCATTGCAAACGGATGCACTACCTAAACCTAAAGTGGACTTAATGCT
6831	ATTTTAGCTAATGAAAAGTGTGGACTAGTACTCTAAATGTCCTAAATCATANGT
476	ATTTTAGCTAATGAAAAGTGTGGACTAGTACTCTAAATGTCCTAAATGATGTTGT

		Query	Match	Similarity	Length	
		Best Local Matches	Conservative	95.8%	673;	Pred. No. 2.8e-47;
		Mismatches	Indels	4;	Gaps	
Db	416	TTTGTAGTAAATAAGAAGAAAAATCTGCAATTCCGGCCAAATTGTATTTTA	357			
Qy	6951	TTCGATTAAATGCTATTCTGTAATTTGGAAATCAGTGGCTTATCATGTATATCG	7010			
Db	356	TGGCGTTAAATGCTATTCTGTAATTTGGAAATCAGTGGCTTATCATGTATATCA	297			
Qy	7011	TGTACTTAAATGPTTAACTACCTGGTATTGTTAATATAAGAACAGATCA	7070			
Db	296	TGACTTAAATGPTTAACTACCTGGTATTGTTAATATAAGAACAGATCA	237			
Qy	7071	TATTTTG-T-GTGTGTATAAGCCTCTAAATAGCATTGAACTTGAAGCTG	7128			
Db	236	TATTTAAGAGATGTTAACTACCTGGTATTGTTAATATAAGAACAGATCA	177			
Qy	7129	ATACTACATTAAACATTACATCCAAAGGAGACTTATTGTCATACACACA	7188			
Db	176	ATACTACATTAAACATTACATCCAAAGGAGACTTATTGTCATACACACA	117			
Qy	7189	TTTAAATAATTAAATTGCTCTAAATAATTAAATTAAATTGTCAGCTGTGGCTACAG	7248			
Db	116	TTTAAATAATTAAATTGCTGTAAATTAAACAACTACTCTGAGCTCTGTGGCTACAG	57			
Qy	7249	TGTTATGTTGCTGTAAAGATAAGATAAGTGTCAA	7281			
Db	56	TGTTATGTTGCTGTAAAGATAAGTGTCAA	24			
		RESULT 15				
LOCUS	AG120959	673 bp	DNA	linear	GSS 04-NOV-2001	
DEFINITION	Pan troglodytes DNA, clone: PTB-129N20.F, genomic survey sequence.					
ACCESSION	AG120959	1	GI:16650124			
VERSION	AG120959.1					
KEYWORDS	Pan troglodytes (chimpanzee)					
SOURCE	Pan troglodytes Eukaryota; Mecazoaa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.					
ORGANISM	1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.					
REFERENCE	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpbees@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170) Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.					
AUTHORS	2 (bases 1 to 673) Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.					
JOURNAL	Direct Submission Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpbees@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170) Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.					
PRIMERS	Sequencing: -21M13 LIBRARY					
	vector	:	pKS145			
	R. Site 1	:	Saci			
	R. Site 2	:	Saci			
	Location/Qualifiers					
	1. .673					
FEATURES	Source /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9588" /clone="PTB-129N20.F" /sex="male" /cell_type="lymphoblast" /clone_id="PTB Chimpanzee Male BAC Library"					
BASE COUNT	238	a	127	c	120	g 188 t
ORIGIN						